(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 May 2001 (17.05.2001)

PCT

(10) International Publication Number
WO 01/34805 A2

(51) International Patent Classification[7]: C12N 15/12, 15/63, C07K 14/705, 16/28, G01N 33/68

(21) International Application Number: PCT/US00/31077

(22) International Filing Date:
10 November 2000 (10.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/438,997 12 November 1999 (12.11.1999) US

(71) Applicant: **ABBOTT LABORATORIES** [US/US]; Dept. 377/ AP6D-2, 100 Abbott Park Road, Abbott Park, IL 60064-6050 (US).

(72) Inventors: **MASTERS, Jeffrey, N.**; 1328 Dunleer Drive, Mundelein, IL 60060 (US). **VOS, Melissa, H.**; 5232 W. Winona Street, Chicago, IL 60630 (US).

(74) Agents: **CASUTO, Dianne** et al.; Dept. 377, AP6D-2, 100 Abbott Park Road, Abbott Park, IL 60064-6050 (US).

(81) Designated States *(national)*: CA, JP, MX.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/34805 A2

(54) Title: HUMAN VANILLOID RECEPTOR GENE

(57) Abstract: The present invention provides an isolated or purified polynucleotide that encodes human vanilloid receptor. Isoforms of human vanilloid receptor are also disclosed. The invention also provides methods of making recombinant human vanilloid receptor using the polynucleotides and host cells transformed with the polynucleotides.

HUMAN VANILLOID RECEPTOR GENE

RELATED APPLICATIONS

This application claims priority to U.S. Application Serial Number 09/191,139 filed November 13, 1998.

TECHNICAL FIELD

The invention relates generally to polynucleotide sequences and polypeptide sequences encoded therefrom, more specifically, to vanilloid receptor genes and polypeptides encoded therefrom as well as methods which utilize these polypeptides for identifying compounds which modulate vanilloid receptors in human tissues.

BACKGROUND OF THE INVENTION

Vanilloid receptors are a class of ligand-gated ion channels defined by the natural ligands capsaicin, the active ingredient of hot peppers from plants of the genus *Capsium* and resiniferitoxin (RTX), an ultrapotent capsaicin analog found in the latex of *Euphorbia resinifera* (Holzer, Pharmacol. Rev. 43:143-201 [1991]). These receptors are involved in a variety of physiological processes including nociception, inflammation, regulation of body temperature, cardiovascular and bronchial systems, reflex bladder function and gastric mucosal defense mechanisms (Capsaicin in the study of pain, Wood ed., 1993, Academic Press).

The vanilloid receptor has been characterized as a cation permeable ion channel with the permeability of di- and mono-valent cations being $Ca^{2+} > Mg^{2+} > K^{+} > Na^{+}$ (Bevan and Szolcsanyl, Trends Pharmacol. 11:330-333 [1990]). Activation of neuronal vanilloid receptors by capsaicin results in initial excitation resulting in pain perception while prolonged exposure results in analgesic effects most likely through a desensitization process (Szallasi, Gen. Pharmac. 25:223-243 [1994]). This biphasic response is characteristic for the other physiological responses to capsaicin, first described for thermoregulation in the hypothalamus (Jancsó-Gábor et al., J. Physiol. 208:449-459 [1970]). The rat vanilloid receptor VR1 has been cloned (Caterina *et al.*, Nature **389:**816-824 [1997]).

Because of capsaicin's role in physiological processes, it would be useful to identify compounds, which modulate the activity of a human vanilloid receptor and/or its analogs. However, efforts to identify such compounds have been hampered by the lack of readily available human vanilloid receptors or cell lines expressing the human vanilloid receptor gene for use in screening assays. Although the rat vanilloid receptor is thought to be expressed exclusively in sensory neurons, human sensory neuron tissue is extremely difficult to obtain. Furthermore, no cell lines have been reported which endogenously express the human vanilloid receptor gene. Thus, there is a need for a simple, easy and cost effective means to obtain large quantities of human vanilloid receptors and/or a cell line which expresses such receptors.

The present invention solves this problem by providing reagents, such as human vanilloid receptors, polynucleotides (and polymorphic variants thereof) which encode for human vanilloid receptors and recombinant expression systems for large-scale production of said receptors. The invention also provides methodologies, such as assays, for identifying compounds which modulate the activity of human vanilloid receptors. Thus, the invention provides high throughput screening assays to identify new vanilloid receptor ligands for the treatment of various disease states including neuropathic pain, inflammation, arthritis, rhinitis, pruritus, bladder dysfunction, cluster headache, wound healing and psoriasis.

SUMMARY OF THE INVENTION

The present invention provides an isolated or purified polynucleotide comprising a nucleotide sequence which encodes a human vanilloid receptor and fragments or complements thereof. Preferably, the nucleotide sequence is SEQ ID NO:1 or fragments thereof. More preferably, the nucleotide sequence is SEQ ID NO:1 from about nucleotide position 435 to about nucleotide position 3050. The invention further provides a polynucleotide comprising a nucleotide sequence which encodes a human vanilloid receptor having the sequence of SEQ ID NO:3.

In another aspect, the polynucleotide can be produced by recombinant techniques. A recombinant molecule comprises a nucleotide sequence that encodes a human vanilloid receptor and is contained within an expression vector. The expression vector may be either a prokaryotic or a eukaryotic vector. Preferred expression vectors are pCIneo and pACSG2. In a more preferred embodiment, the nucleotide sequence which encodes a human vanilloid receptor has the sequence SEQ ID NO:1 from about nucleotide position 435 to about nucleotide position 3050.

The present invention further provides a host cell transformed with said vector. The host cell is either a prokaryotic or eukaryotic cell.

The present invention also provides a polypeptide of a human vanilloid receptor or fragments thereof. In a preferred embodiment, the polypeptide has the amino acid sequence SEQ ID NO:3. The polypeptide can be produced by recombinant technology and provided in purified form.

In another aspect, the invention provides a method for producing a polypeptide which contains at least one human vanilloid receptor epitope, wherein the method comprises incubating host cells transformed with an expression vector comprising a nucleotide sequence which encodes a human vanilloid receptor. Preferably, the expression vector comprises a nucleotide sequence having the sequence SEQ ID NO:1 and fragments and complements thereof. More preferably, the nucleotide sequence has the sequence SEQ ID NO:1 from about

nucleotide position 435 to about nucleotide position 3050. Even more preferably, the nucleotide sequence encodes a human vanilloid receptor having sequence SEQ ID NO:3.

In another aspect, the invention provides a method for identifying compounds that modulate vanilloid receptor activity, comprising the steps of: (a) providing a host cell that expresses the vanilloid receptor polypeptide; (b) mixing a test compound with the cell; and (c) measuring either (i) the effect of the test compound on the cell expressing the receptor, or (ii) the binding of the test compound to the cell or to the receptor. The host cell of the method is either a prokaryotic or eukaryotic cell. Preferably in the method, the measurement of step (c)(ii) is performed by measuring a signal generated by a signal-generating compound or by measuring a signal generated by a radiolabeled ion, a fluorescent probe or an electrical current.

In yet another aspect, the invention provides a method for identifying a cytoprotective compound, comprising the steps of: (a) providing a cell that expresses a vanilloid receptor polypeptide or fragment thereof; (b) combining a test compound with the cell; and (c) monitoring the cell or cellular function for an indication of cytotoxicity. The host cell of the method is either a prokaryotic or eukaryotic cell. Preferably, the method comprises providing a cell which has an expression vector comprising a polynucleotide having the nucleotide sequence SEQ ID NO:1 from about nucleotide position 435 to about nucleotide position 3050 operably linked to control sequences that direct the transcription of the polynucleotide whereby the polynucleotide is expressed in a host cell. More preferably, one of the control sequences comprises an inducible promotor . Even more preferably, the cell is maintained in the presence of a substance which minimizes or blocks a cytotoxic effect on the cell.

In yet another aspect, the invention provides a method of treating an individual having a condition associated with vanilloid receptor modulation, comprising administering to the individual an effective amount of a compound that controls the gene expression of vanilloid receptor, in a pharmaceutically acceptable excipient.

In yet another embodiment, the invention provides a monoclonal antibody or a polyclonal antibody which specifically binds to human vanilloid receptor having amino acid sequence SEQ ID NO:3 or fragments thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows the alignment (default parameters of the FRAMEALIGN Program, Wisconsin Sequence Analysis Package, Version 9, Genetics Computer Group, Madison, WI) between the consensus sequence (upper line) of the overlapping human Incyte ESTs 1427917 (nt 1-227) and 3460342 (nt 32-270) with the corresponding amino acid sequence of the rat vanilloid receptor (bottom line, Caterina et al., 1997, *supra*). In this Figure, the upper line

corresponds to nucleotides 1696-1966 of SEQ ID NO:7. The bottom line corresponds to amino acid residues 500-589 of SEQ ID NO:4.

FIG. 2 shows the alignment (default parameters of the GAP Program, Wisconsin Sequence Analysis Package, Version 9, Genetics Computer Group, Madison, WI) between cDNA sequences of the human vanilloid receptor (hVR1, top line, SEQ ID NO:7) and rat vanilloid receptor 1 (rVR1, bottom line, SEQ ID NO:2, Caterina et al., 1997, *supra*). Vertical lines between the two sequences indicate identical nucleotides at those positions. Methionine initiation codons (ATG) and stop codons (TGA in top strand and TAA in bottom strand) are boxed. The DNA sequence identity of the rat cDNA from nucleotides (nt) 44-2730 and the human cDNA from nt 163-2874 is 82%.

FIG. 3 shows the multiple alignment (default parameters of the Pileup and Pretty Programs, Wisconsin Sequence Analysis Package, Version 10, Genetics Computer Group, Madison, WI) of the amino acid sequences of the hVR1 (SEQ ID NO:8), rVH1 (SEQ ID NO:4, Caterina et al., 1997, *supra*), human vanilloid receptor-like protein (hVR2, Caterina et al., Nature 398:436-441 1999, SEQ ID NO:15) and human vanilloid receptor 3 (hVR3, SEQ ID NO:3). The consensus sequence identifies any identical amino acid position shared by these 4 proteins. Boxed regions indicate the ankaryn repeats (position 239-270, 294-317 and 370-403), potential transmembrane domains (positions 471-493, 519-540, 555-574, 579-597, 621-640 and 703-730) and the poor-loop region (position 671-691).

FIG. 4 shows the polymorphic regions of the human vanilloid receptor determined by direct sequencing of the PCR product of human small intestine RNA. The Sequencher™ chromatogram tracings (Sequencher™ Version 3.0, Gene Codes Corp., Ann Arbor, MI) are shown with arrows identifying the double peaks consistent with polymorphic positions. Nt position 1605 contains either a C or a T while nt position 1952 contains an A or a G.

FIG. 5 shows the GAP analysis of hVR1 (bottom sequence, positions 301-3410 of SEQ ID NO:7) and hVR3 (top sequence, positions 1-3055 of SEQ ID NO:1) DNA sequences.

FIG. 6 shows the GAP analysis of the derived amino acid sequences of hVR1 (bottom sequence) and hVR3 (top sequence), SEQ ID NOs:8 and 3 respectively.

FIG. 7 shows a graphical representation of expression of hVR1 and hVR3 by quantitative RT-PCR (ABI Prism 7700) of total RNA isolated from human adrenal gland (lane 1), brain (lane 2), cerebellum (lane 3), fetal brain (lane 4), fetal liver (lane 5), heart (lane 6), kidney (lane 7), liver (lane 8), lung (lane 9), mammary gland (lane 10), pancreas (lane 11), placenta (lane 12), prostate (lane 13), salivary gland (lane 14), skeletal muscle (lane 15), small intestine (lane 16), spleen (lane 17), stomach (lane 18), testes (lane 19), thymus (lane 20), trachea (lane 21), uterus (lane 22), DRG (lane 23), bladder (lane 24) and HEK293 cells (lane 25) using primers specific for hVR1 and hVR3. The hatched bars represent samples from an additional experiment.

FIG. 8 shows the nucleotide sequence (SEQ ID NO:1) of human vanilloid receptor 3 and deduced amino acid sequence (SEQ ID NO:3).

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides isolated and purified polynucleotides that encode a human vanilloid receptor, fragments thereof, expression vectors containing those polynucleotides, host cells transformed with those expression vectors, a process for making a human vanilloid receptor using those polynucleotides and vectors, and isolated and purified recombinant human vanilloid receptor and polypeptide fragments thereof.

Portions of the nucleic acid sequences disclosed herein are useful as primers for the reverse transcription of RNA or for the amplification of cDNA; or as probes to determine the presence of certain cDNA sequences in test samples. Also disclosed are nucleic acid sequences which permit the production of encoded polypeptide sequences which are useful as standards or reagents in diagnostic immunoassays, targets for pharmaceutical screening assays and/or as components or target sites for various therapies. Isolation of sequences from other portions of the vanilloid receptor gene can be accomplished by utilizing probes or PCR primers derived from these nucleic acid sequences, thus allowing additional probes and polypeptides of the genome of interest to be established.

The present invention also provides methods for assaying a test sample for products of a human vanilloid receptor gene, which comprises making cDNA from mRNA in the test sample, and detecting the cDNA as an indication of the presence of a human vanilloid receptor gene. The method may include an amplification step, wherein portions of the cDNA corresponding to the gene or fragment thereof is amplified. Methods also are provided for assaying for the translation products of mRNAs. Test samples which may be assayed by the methods provided herein include tissues, cells, body fluids and secretions. The present invention also provides reagents such as oligonucleotide primers and polypeptides which are useful in performing these methods. For example, the invention provides monoclonal and polyclonal antibodies directed against at least one epitope contained within the polypeptide sequences of the invention which are useful for diagnostic tests and for screening for diseases or conditions associated with abnormal vanilloid receptor production.

Although the physiological manifestations of abnormal vanilloid receptor expression are as yet unknown in humans or other mammals, we postulate that the vanilloid receptor may play a pathological role resulting from its abnormal expression. For example, Caterina et al. have shown that HEK293 cells transfected with the vanilloid receptor are killed within several hours of continuous exposure to capsaicin. Therefore, it is reasonable to postulate that the presence of

vanilloid receptor in certain body fluids where it is not normally found may be indicative of a disease state, the further progression of which could be monitored by assaying for vanilloid receptor in such fluids. A similar role is seen for myelin basic protein (MBP) which in the normal physiological state is a membrane bound protein and therefore not found in body fluids, but in disease states such as multiple sclerosis, it is released into cerebral spinal fluid. Furthermore, the presence of a polynucleotide or fragment thereof which encodes vanilloid receptor in tissues or body fluids where it is unexpected, may also be indicative of a disease condition, in the case, for example, where the disease was manifest by cellular degeneration.

Thus, the reagents and methods described herein may enable the identification of certain markers as indicative of abnormal vanilloid receptor expression and the information obtained therefrom may aid in the diagnosis, staging, monitoring, prognosis and/or therapy of diseases or conditions which may be associated with such expression. Test methods include, for example, probe assays which utilize the sequence(s) provided herein and which also may utilize nucleic acid amplification methods such as the polymerase chain reaction (PCR), the ligase chain reaction (LCR); and hybridization. In addition, the nucleotide sequences provided herein contain open reading frames from which an immunogenic epitope may be found. Preferably, such an epitope is unique to the disease state or condition associated with the vanilloid receptor gene. The uniqueness of the epitope may be determined by its immunological reactivity with the polypeptide product encoded by such gene, and lack of immunological reactivity with tissue(s) from non-diseased patients. Methods for determining immunological reactivity are well-known and include but are not limited to, for example, radioimmunoassay (RIA), enzyme-linked immunosorbent assay (ELISA), hemagglutination (HA), fluorescence polarization immunoassay (FPIA); chemiluminescent immunoassay (CLIA), and others; several examples of suitable methods are described herein.

Definitions

All patents, patent applications and publications cited herein, whether *supra* or *infra*, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a", "an", and "the" include plural references unless the content clearly dictates otherwise.

Unless otherwise stated, the following terms shall have the following meanings:

"Purified product" refers to a preparation of the product, which has been isolated from the cellular constituents with which the product is normally associated, and from other types of cells, which may be present in the sample of interest.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same

polynucleotide or DNA or polypeptide, which is separated from some or all of the coexisting materials in the natural system, is isolated. Such a polynucleotide could be part of a vector and/or such a polynucleotide or polypeptide could be part of a composition, and still be isolated in that the vector or composition is not part of its natural environment.

The term "polynucleotide" as used herein means a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, the term includes double- and single-stranded DNA, as well as, double- and single-stranded RNA. It also includes modifications, such as methylation or capping, and unmodified forms of the polynucleotide.

A "vanilloid receptor variant" refers to an isolated vanilloid receptor polynucleotide sequence having at least 83%, more preferably, at least 90% and even more preferably, at least 95% global sequence identity over a length of a vanilloid receptor polynucleotide, to vanilloid receptor polynucleotides disclosed herein. "Percent identity" is determined using the default parameters of the GAP program, Wisconsin Sequence Analysis Package, Version 9, Genetics Computer Group, Madison, WI).

A "polynucleotide fragment derived from" a designated sequence refers to a polynucleotide sequence which is comprised of a sequence of approximately at least about 6 nucleotides, is preferably at least about 8 nucleotides, is more preferably at least about 10, is more preferably at least about 12 nucleotides, is more preferably at least about 15 and even more preferably is at least about 20 nucleotides corresponding, i.e., identical to or complementary to, a region of the designated nucleotide sequence. The sequence may be complementary to or identical to a sequence which is unique to a particular polynucleotide sequence as determined by techniques known in the art. Comparisons to sequences in databanks, for example, can be used as a method to surmise the uniqueness of a designated sequence. Regions from which sequences may be derived include but are not limited to regions encoding specific epitopes, as well as non-translated and/or non-transcribed regions.

The derived polynucleotide will not necessarily be derived physically from the nucleotide sequence of interest under study, but may be generated in any manner, including but not limited to chemical synthesis, replication, reverse transcription or transcription, which is based on the information provided by the sequence of bases in the region(s) from which the polynucleotide is derived; as such, it may represent either a sense or an antisense orientation of the original polynucleotide. In addition, combinations of regions corresponding to that of the designated sequence may be modified in ways known in the art to be consistent with an intended use.

When referring to a nucleic acid fragment, such a fragment is considered to "selectively hybridize" or to "selectively bind" to a polynucleotide or variants thereof disclosed herein, if, within the linear range of detection, the hybridization results in a

stronger signal relative to the signal that results from hybridization of the fragment to an equal amount of a second polynucleotide. A signal which is "stronger" than another is one which is measurable over the other by the particular method of detection. Methods for hybridizing and detecting polynucleotides are well known to those of ordinary skill in the art.

Also, when referring to a nucleic acid fragment, such a fragment is considered to hybridize under selective hybridization conditions if it selectively hybridizes under (i) typical hybridization and wash conditions, such as those described, for example, in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, second edition, (1989), Cold Spring Harbor, N.Y. and Nucleic Acid Hybridization: A Practical Approach, editors B.D. Hames and S.J. Higgins, (1985) Oxford; Washington D.C.; IRL Press), where preferred hybridization conditions are those of lesser stringency and more preferred, higher stringency; or (ii) standard PCR conditions (Saiki, R.K. *et al.* (1988) Science. 239:487-491) or "touch-down" PCR conditions (Roux, K.H., (1994), Biotechiques, 16:812-814).

"A sequence corresponding to a cDNA" means that the sequence contains a polynucleotide sequence that is identical to or complementary to a sequence in the designated DNA. The degree (or "percent") of identity or complementarity to the cDNA will be approximately 50% or greater, will preferably be at least about 70% or greater, and more preferably will be at least about 90% or greater. The sequence that corresponds will be at least about 50 nucleotides in length, will preferably be about 60 nucleotides in length, and more preferably, will be at least about 70 nucleotides in length. The correspondence between the gene or gene fragment of interest and the cDNA can be determined by methods known in the art, and include, for example, a direct comparison of the sequenced material with the cDNAs described, or hybridization and digestion with single strand nucleases, followed by size determination of the digested fragments.

"Purified polynucleotide" refers to a polynucleotide of interest or fragment thereof which is essentially free, i.e., contains less than about 50%, preferably less than about 70%, and more preferably, less than about 90% of the protein with which the polynucleotide is naturally associated. Techniques for purifying polynucleotides of interest are well-known in the art and include, for example, disruption of the cell containing the polynucleotide with a chaotropic agent and separation of the polynucleotide(s) and proteins by ion-exchange chromatography, affinity chromatography and sedimentation according to density. Thus, "purified polypeptide" means a polypeptide of interest or fragment thereof which is essentially free, that is, contains less than about 50%, preferably less than about 70%, and more preferably, less than about 90%

of cellular components with which the polypeptide of interest is naturally associated. Methods for purifying are known in the art.

The term "probe" denotes a defined nucleic acid segment (or nucleotide analog segment, i.e., peptide nucleic acid analog (PNA) or morpholino analog (MA) which can be used to identify specific DNA or RNA present in samples bearing the complementary sequence.

The term "primer" denotes a specific oligonucleotide sequence complementary to a target nucleotide sequence and used to hybridize to the target nucleotide sequence and serve as an initiation point for nucleotide polymerization catalyzed by either DNA polymerase or reverse transcriptase.

"Polypeptide" as used herein indicates a molecular chain of amino acids and does not refer to a specific length of the product. Thus, peptides, oligopeptides and proteins are included within the definition of polypeptide. This term, however, is not intended to refer to post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like.

A "recombinant polypeptide" as used herein means at least a polypeptide which by virtue of its origin or manipulation is not associated with all or a portion of the polypeptide with which it is associated in nature and/or is linked to a polypeptide other than that to which it is linked in nature. A recombinant or derived polypeptide is not necessarily translated from a designated nucleic acid sequence. It also may be generated in any manner, including chemical synthesis or expression of a recombinant expression system.

The term "synthetic peptide" as used herein means a polymeric form of amino acids of any length, which may be chemically synthesized by methods well-known to those of ordinary skill in the art. These synthetic peptides are useful in various applications.

A vanilloid receptor polypeptide, as used herein, refers to polypeptide having at least 87%, more preferably at least 90%, and even more preferably at least 95% global sequence identity over a length of a vanilloid receptor polypeptide, to vanilloid receptor polypeptides disclosed herein. A most preferred vanilloid receptor polypeptide is SEQ ID NO:3. Two other preferred vanilloid receptor polypeptides have essentially identical sequences to SEQ ID NO:3 with the exception that in one preferred polypeptide, at residue 469, threonine is replaced by isoleucine and in the second preferred polypeptide, at residue 586, isoleucine is replaced with valine.

A "polypeptide or amino acid sequence derived from" a designated nucleic acid sequence refers to a polypeptide having an amino acid sequence identical to that of a polypeptide encoded in the sequence or a portion thereof wherein the portion consists of at least 3 to 5 amino acids, and more preferably at least 8 to 10 amino acids, and even more preferably 15 to 20 amino acids, or which is immunologically identifiable with a polypeptide encoded in the sequence.

As used herein, the term "identity" refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. The term "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" can be determined between the compared polypeptide sequences. Further, the polypeptide or amino acid sequence may preferably have at least 90% similarity, more preferably about 95% similarity and most preferably about 98% similarity to a polypeptide or amino acid sequence of a human vanilloid receptor.

The percent identity of two sequences, whether nucleic acid or peptide sequences, is the number of exact matches between two aligned sequences divided by the length of the shorter sequences and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981). This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Schwartz, R.M., and Dayhoff, M.O. Matrices for detecting distant relationships, (in) Atlas of Protein Sequence and Structure, 5 suppl.3:353-358, (Nat. Biomed. Res. Found., Washington D.C.), 1978. and normalized by Gribskov, Nucl. Acids Res. 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their GAP utility application. The default parameters for this method are described in the Wisconsin Sequence analysis Package, Program Manual, Version 9 (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

Other techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded therein, and comparing this to a second amino acid sequence.

"Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refer to cells which can be, or have been, used as recipients for recombinant vector or other transferred DNA, and include the original progeny of the original cell which has been transfected.

As used herein "replicon" means any genetic element, such as a plasmid, a chromosome or a virus, that behaves as an autonomous unit of polynucleotide replication within a cell.

A "vector" is a replicon in which another polynucleotide segment is attached, such as to bring about the replication and/or expression of the attached segment.

The term "control sequence" refers to polynucleotide sequences which are necessary to effect the expression of coding sequences to which they are ligated. The nature of such control sequences differs depending upon the host organism. In prokaryotes, such control sequences generally include promoter, ribosomal binding site and terminators; in eukaryotes, such control sequences generally include promoters, terminators and, in some instances, enhancers. The term "control sequence" thus is intended to include at a minimum all components whose presence is necessary for expression, and also may include additional components whose presence is advantageous, for example, leader sequences.

"Operably linked" refers to a situation wherein the components described are in a relationship permitting them to function in their intended manner. Thus, for example, a control sequence "operably linked" to a coding sequence is ligated in such a manner that expression of the coding sequence is achieved under conditions compatible with the control sequences.

The term "open reading frame" or "ORF" refers to a region of a polynucleotide sequence which encodes a polypeptide; this region may represent a portion of a coding sequence or a total coding sequence.

A "coding sequence" is a polynucleotide sequence which is transcribed into mRNA and translated into a polypeptide when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the 5' -terminus and a translation stop codon at the 3' -terminus. A coding sequence can include, but is not limited to, mRNA, cDNA, and recombinant polynucleotide sequences.

The term "sense strand" or "plus strand" (or "+") as used herein denotes a nucleic acid that contains the sequence that encodes the polypeptide. The term "antisense strand" or "minus strand" (or "-") denotes a nucleic acid that contains a sequence that is complementary to that of the "plus" strand.

An "Expressed Sequence Tag" or "EST" refers to the partial sequence of a cDNA insert which has been made by reverse transcription of mRNA extracted from a tissue, followed by insertion into a vector.

A "transcript image" refers to a table or list giving the quantitative distribution of ESTs in a library and represents the genes active in the tissue from which the library was made.

The term "immunologically identifiable with/as" refers to the presence of epitope(s) and polypeptide(s) which also are present in and are unique to the designated polypeptide(s). Immunological identity may be determined by antibody binding and/or competition in binding. These techniques are known to those of ordinary skill in the art and also are described herein. The uniqueness of an epitope also can be surmised by computer searches of known data banks, such as GenBank, for the polynucleotide sequences which encode the epitope, and by amino acid sequence comparisons with other known proteins.

As used herein, "epitope" means an antigenic determinant of a polypeptide. Conceivably, an epitope can comprise three amino acids in a spatial conformation which is unique to the epitope. Generally, an epitope consists of at least five such amino acids, and more usually, it consists of at least eight to ten amino acids. Methods of examining spatial conformation are known in the art and include, for example, x-ray crystallography and two-dimensional nuclear magnetic resonance.

A "conformational epitope" is an epitope that is comprised of specific juxtaposition of amino acids in an immunologically recognizable structure, such amino acids being present on the same polypeptide in a contiguous or non-contiguous order or present on different polypeptides.

A polypeptide is "immunologically reactive" with an antibody when it binds to an antibody due to antibody recognition of a specific epitope contained within the polypeptide. Immunological reactivity may be determined by antibody binding, more particularly by the kinetics of antibody binding, and/or by competition in binding using as competitor(s) a known polypeptide(s) containing an epitope against which the antibody is directed. The methods for determining whether a polypeptide is immunologically reactive with an antibody are known in the art.

As used herein, the term "immunogenic polypeptide containing an epitope of interest" means naturally occurring polypeptides of interest or fragments thereof, as well as polypeptides prepared by other means, for example, by chemical synthesis or the expression of the polypeptide in a recombinant organism.

The terms "transformation" refers to the insertion of an exogenous polynucleotide into a prokaryotic or yeast host cell, irrespective of the method used for the insertion. Generally, the term "transfection" is used with respect to insertion of an exogenous polynucleotide into a eukaryotic host cell. The processes for achieving transformation and/or transfection are well known to those of ordinary skill in the art and include such techniques as direct uptake, transduction, f-mating and electroporation. The exogenous polynucleotide may be maintained as a non-integrated vector, for example, a plasmid, or alternatively, may be integrated into the host genome.

"Treatment" refers to prophylaxis and/or therapy.

The term "individual" as used herein refers to vertebrates, particularly members of the mammalian species and includes but is not limited to domestic animals, sports animals, primates and humans; more particularly the term refers to humans.

The term "test sample" refers to a component of an individual's body which is the source of the analyte (also referred to "target" or "marker"). These components include antibodies and antigens and are well known in the art. These test samples include biological samples which can be tested by the methods of the present invention described herein and

include human and animal body fluids such as whole blood, serum, plasma, cerebrospinal fluid, urine, lymph fluids, and various external secretions of the respiratory, intestinal and genitor-urinary tracts, tears, saliva, milk, white blood cells, myelomas and the like; biological fluids such as cell culture supernatants; fixed tissue specimens; and fixed cell specimens.

"PNA" denotes a "peptide nucleic acid analog" which may be utilized in a procedure such as an assay described herein to determine the presence of a target. "MA" denotes a "morpholino analog" which may be utilized in a procedure such as an assay described herein to determine the presence of a target. See, for example, U.S. Patent No. 5,378,841, which is incorporated herein by reference. PNAs are neutrally charged moieties which can be directed against RNA targets or DNA. PNA probes used in assays in place of, for example, the DNA probes of the present invention, offer advantages not achievable when DNA probes are used. These advantages include manufacturability, large scale labeling, reproducibility, stability, insensitivity to changes in ionic strength and resistance to enzymatic degradation which is present in methods utilizing DNA or RNA. These PNAs can be labeled with such signal generating compounds as fluorescein, radionucleotides, chemiluminescent compounds, and the like. PNAs or other nucleic acid analogs such as MAs thus can be used in assay methods in place of DNA or RNA. Although assays are described herein utilizing DNA probes, it is within the scope of the routineer that PNAs or MAs can be substituted for RNA or DNA with appropriate changes if and as needed in assay reagents.

"Analyte," as used herein, is the substance to be detected which may be present in the test sample. The analyte can be any substance for which there exists a naturally occurring specific binding member (such as, an antibody), or for which a specific binding member can be prepared. Thus, an analyte is a substance that can bind to one or more specific binding members in an assay. "Analyte" also includes any antigenic substances, haptens, antibodies, and combinations thereof. As a member of a specific binding pair, the analyte can be detected by means of naturally occurring specific binding partners (pairs) such as the use of intrinsic factor protein as a member of a specific binding pair for the determination of Vitamin B12, the use of folate-binding protein to determine folic acid, or the use of a lectin as a member of a specific binding pair for the determination of a carbohydrate. The analyte can include a protein, a peptide, an amino acid, a nucleotide target, and the like.

The present invention provides assays which utilize specific binding members. A "specific binding member," as used herein, is a member of a specific binding pair. That is, two different molecules where one of the molecules through chemical or physical means specifically binds to the second molecule. Therefore, in addition to antigen and antibody specific binding pairs of common immunoassays, other specific binding pairs can include biotin and avidin, carbohydrates and lectins, complementary nucleotide sequences, effector and receptor molecules, cofactors and enzymes, enzyme inhibitors and enzymes, and the like.

Furthermore, specific binding pairs can include members that are analogs of the original specific binding members, for example, an analyte-analog. Immunoreactive specific binding members include antigens, antigen fragments, antibodies and antibody fragments, both monoclonal and polyclonal, and complexes thereof, including those formed by recombinant DNA molecules.

The term "hapten," as used herein, refers to a partial antigen or non-protein binding member which is capable of binding to an antibody, but which is not capable of eliciting antibody formation unless coupled to a carrier protein.

A "capture reagent," as used herein, refers to an unlabeled specific binding member which is specific either for the analyte as in a sandwich assay, for the indicator reagent or analyte as in a competitive assay, or for an ancillary specific binding member, which itself is specific for the analyte, as in an indirect assay. The capture reagent can be directly or indirectly bound to a solid phase material before the performance of the assay or during the performance of the assay, thereby enabling the separation of immobilized complexes from the test sample.

The "indicator reagent" comprises a "signal-generating compound" ("label") which is capable of generating and generates a measurable signal detectable by external means, conjugated ("attached") to a specific binding member. The indicator reagent can be a member of any specific binding pair including hapten-anti-hapten systems such as biotin or anti-biotin, avidin or biotin, a carbohydrate or a lectin, a complementary nucleotide sequence, an effector or a receptor molecule, an enzyme cofactor and an enzyme, an enzyme inhibitor or an enzyme, and the like. An immunoreactive specific binding member can be an antibody, an antigen, or an antibody/antigen complex that is capable of binding either to polypeptide of interest as in a sandwich assay, to the capture reagent as in a competitive assay, or to the ancillary specific binding member as in an indirect assay. When describing probes and probe assays, the term "reporter molecule" may be used. A reporter molecule comprises a signal generating compound as described hereinabove conjugated to a specific binding member of a specific binding pair, such as carbazol or adamantane.

The various "signal-generating compounds" (labels) contemplated include chromogens, catalysts such as enzymes, luminescent compounds such as fluorescein and rhodamine, chemiluminescent compounds such as dioxetanes, acridiniums, phenanthridiniums and luminol, radioactive elements, and direct visual labels. Examples of enzymes include alkaline phosphatase, horseradish peroxidase, beta-galactosidase, and the like. The selection of a particular label is not critical, but it will be capable of producing a signal either by itself or in conjunction with one or more additional substances.

"Solid phases" ("solid supports") are known to those in the art and include the walls of wells of a reaction tray, test tubes, polystyrene beads, magnetic beads, nitrocellulose strips, membranes, microparticles such as latex particles, sheep (or other animal) red blood cells, and

Duracytes® (red blood cells "fixed" by pyruvic aldehyde and formaldehyde, available from Abbott Laboratories, Abbott Park, IL) and others. The "solid phase" is not critical and can be selected by one skilled in the art. Thus, latex particles, microparticles, magnetic or non-magnetic beads, membranes, plastic tubes, walls of microtiter wells, glass or silicon chips, sheep (or other suitable animal's) red blood cells and Duracytes® are all suitable examples. Suitable methods for immobilizing peptides on solid phases include ionic, hydrophobic, covalent interactions and the like. A "solid phase", as used herein, refers to any material which is insoluble, or can be made insoluble by a subsequent reaction. The solid phase can be chosen for its intrinsic ability to attract and immobilize the capture reagent. Alternatively, the solid phase can retain an additional receptor which has the ability to attract and immobilize the capture reagent. The additional receptor can include a charged substance that is oppositely charged with respect to the capture reagent itself or to a charged substance conjugated to the capture reagent. As yet another alternative, the receptor molecule can be any specific binding member which is immobilized upon (attached to) the solid phase and which has the ability to immobilize the capture reagent through a specific binding reaction. The receptor molecule enables the indirect binding of the capture reagent to a solid phase material before the performance of the assay or during the performance of the assay. The solid phase thus can be a plastic, derivatized plastic, magnetic or non-magnetic metal, glass or silicon surface of a test tube, microtiter well, sheet, bead, microparticle, chip, sheep (or other suitable animal's) red blood cells, Duracytes® and other configurations known to those of ordinary skill in the art.

It is contemplated and within the scope of the present invention that the solid phase also can comprise any suitable porous material with sufficient porosity to allow access by detection antibodies and a suitable surface affinity to bind antigens. Microporous structure generally are preferred, but materials with gel structure in the hydrated state may be used as well. Such useful solid supports include but are not limited to nitrocellulose and nylon. It is contemplated that such porous solid supports described herein preferably are in the form of sheets of thickness from about 0.01 to 0.5 mm, preferably about 0.1mm. The pore size may vary within wide limits, and preferably is from about 0.025 to 15 microns, especially from about 0.15 to 15 microns. The surface of such supports may be activated by chemical processes which cause covalent linkage of the antigen or antibody to the support. The irreversible binding of the antigen or antibody is obtained, however, in general, by adsorption on the porous material by poorly understood hydrophobic forces. Other suitable solid supports are known in the art.

Reagents

The present invention provides reagents such as polynucleotide sequences derived from a human vanilloid receptor gene, polypeptides encoded therein, and antibodies produced from these polypeptides. The present invention also provides reagents such as oligonucleotide

fragments derived from the disclosed polynucleotides and nucleic acid sequences complementary to these polynucleotides. For example, selected vanilloid receptor-derived polynucleotides can be used in the methods described herein for the detection of normal or altered gene expression.

The present invention also provides methods, in particular, recombinant methodologies using polynucleotide sequences disclosed herein, for making human vanilloid receptors in high yield as well as methods to identify compounds which modulate (i.e. activate or repress) the activity of such receptors. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably purified.

Furthermore, the polynucleotides disclosed herein, their complementary sequences or fragments of either can be used in assays to detect, amplify or quantify genes, cDNAs or mRNAs encoding human vanilloid receptor. They also can be used to identify an entire or partial coding region which encodes for a vanilloid receptor polypeptide. They further can be provided in individual containers in the form of a kit for assays, or provided as individual compositions. If provided in a kit for assays, other suitable reagents such as buffers, conjugates and the like may be included.

The polynucleotide(s) may be in the form of mRNA or DNA. Polynucleotides in the form of DNA, cDNA, genomic DNA, and synthetic DNA are within the scope of the present invention. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding (sense) strand or non-coding (antisense) strand. The coding sequence which encodes the polypeptide may be identical to the coding sequence provided herein or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same polypeptide as the DNA provided herein.

This polynucleotide may include only the coding sequence for the polypeptide, or the coding sequence for the polypeptide and additional coding sequence such as a leader or secretory sequence or a proprotein sequence, or the coding sequence for the polypeptide (and optionally additional coding sequence) and non-coding sequence, such as a non-coding sequence 5' and/or 3' of the coding sequence for the polypeptide.

In addition, the invention includes variant polynucleotides containing modifications such as polynucleotide deletions, substitutions or additions; and any polypeptide modification resulting from the variant polynucleotide sequence. A polynucleotide of the present invention also may have a coding sequence which is a naturally occurring allelic variant of the coding sequence provided herein.

In addition, the coding sequence for the polypeptide may be fused in the same reading frame to a polynucleotide sequence which aids in expression and secretion of a polypeptide from a host cell, for example, a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide from the cell. The polypeptide having a leader sequence

is a preprotein and may have the leader sequence cleaved by the host cell to form the form of the polypeptide. The polynucleotides may also encode for a proprotein which is the protein plus additional 5' amino acid residues. A protein having a prosequence is a proprotein and may in some cases be an inactive form of the protein. Once the prosequence is cleaved an active protein remains. Thus, the polynucleotide of the present invention may encode for a protein, or for a protein having a prosequence or for a protein having both a presequence (leader sequence) and a prosequence.

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptide of the present invention. The marker sequence may be a hexa-histidine tag supplied by a pProEx1 (Life Technologies, Gaithersburg, MD) vector to provide for purification of the polypeptide fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein. See, for example, I. Wilson, *et al.*, Cell 37:767 (1984).

It is contemplated that polynucleotides which encode a human vanilloid receptor will be considered to hybridize to the sequences provided herein if there is at least 60%, more preferably at least 70% and even more preferably at least 80%, identity between the polynucleotide and the sequence.

The present invention further provides human vanilloid receptor polypeptides which have the deduced amino acid sequences as provided herein, as well as fragments, analogs and derivatives of such polypeptides. The polypeptides of the present invention may be recombinant polypeptides, natural purified polypeptides or synthetic polypeptides. The polypeptides, fragments, derivatives or analogs of the human vanilloid receptor may be those in which one or more of the amino acid residues is substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code; or it may be one in which one or more of the amino acid residues includes a substituent group; or it may be one in which the polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol); or it may be one in which the additional amino acids are fused to the polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the polypeptide or a proprotein sequence.

Thus, a polypeptide of the present invention may have an amino acid sequence that is identical to that of the naturally occurring polypeptide or that is different by minor variations due to one or more amino acid substitutions. The variation may be a "conservative change" typically in the range of about 1 to 5 amino acids, wherein the substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine or

threonine with serine. In contrast, variations may include nonconservative changes, eg, replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which and how many amino acid residues may be substituted, inserted or deleted without changing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software (DNASTAR Inc., Madison, WI).

A human vanilloid polypeptide is believed to be substantially encoded by or within SEQ ID NO:1 or SEQ ID NO:7. The minimum polypeptide sequence necessary for ligand binding, however, can be determined by routine methods. The sequence, for example, may be truncated at either end by treating an appropriate expression vector with an exonuclease (after cleavage at the 5' or 3' end of the coding sequence) to remove any desired number of base pairs. The resulting coding polynucleotide is then expressed and the sequence determined. In this manner the binding activity may be correlated with the amino acid sequence: a limited series of such experiments (removing progressively greater numbers of base pairs) determines the minimum internal sequence necessary for ligand-binding activity.

The vanilloid receptor polypeptides may be naturally purified products expressed from a high expressing cell line, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture) as described above. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated with mammalian or other eukaryotic carbohydrates or may be non-glycosylated. The polypeptides of the invention may also include an initial methionine amino acid residue. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers or produced by cell-free translation systems using RNAs derived from the DNA constructs of the present invention.

The present invention also provides an antibody produced by using a purified vanilloid receptor gene polypeptide of which at least a portion of the polypeptide is encoded by a vanilloid receptor gene polynucleotide selected from the polynucleotides provided herein. These antibodies may be used in the methods provided herein for the detection of vanilloid receptor polypeptides in test samples. The antibody also may be used for therapeutic purposes, for example, in neutralizing the activity of a vanilloid receptor polypeptide in conditions associated with its altered or abnormal expression. The antibody may also be used to detect an accessory protein or proteins by immunoprecipitation of protein complexes.

Probe Assays

The sequences provided herein may be used to produce probes which can be used in assays for the detection of nucleic acids in test samples. For example, such probes can be used

in Fluorescent In Situ Hybridization (FISH) technology to perform chromosomal analysis, and used to identify vanilloid receptor structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR-generated and/or allele specific oligonucleotide probes, allele specific amplification or by direct sequencing. Probes also can be labeled with radioisotopes, directly- or indirectly- detectable haptens, or fluorescent molecules, and utilized for in situ hybridization studies to evaluate the mRNA expression of the gene comprising the polynucleotide in fixed tissue specimens or cells.

The probes may be designed from conserved nucleotide regions of the polynucleotides of interest or from non-conserved nucleotide regions of the polynucleotide of interest. The design of such probes for optimization in assays is within the skill of the routineer. Generally, nucleic acid probes are developed from non-conserved or unique regions when maximum specificity is desired, and nucleic acid probes are developed from conserved regions when assaying for nucleotide regions that are closely related to, for example, different members of a multigene family or in related species like mouse and man.

The polymerase chain reaction (PCR) is a technique for amplifying a desired nucleic acid sequence (target) contained in a nucleic acid or mixture thereof. In PCR, a pair of primers are employed in excess to hybridize at the outside ends of complementary strands of the target nucleic acid. The primers are each extended by a polymerase using the target nucleic acid as a template. The extension products become target sequences themselves, following dissociation from the original target strand. New primers then hybridize to the target sequences and are extended by a polymerase, and the cycle is repeated to geometrically increase the number of target sequence molecules. PCR is disclosed in U.S. patents 4,683,195 and 4,683,202.

The Ligase Chain Reaction (LCR) is an alternate method for nucleic acid amplification. In LCR, probe pairs are used which include two primary (first and second) and two secondary (third and fourth) probes, all of which are employed in molar excess to a target. The first probe hybridizes to a first segment of the target strand and the second probe hybridizes to a second segment of the target strand, the first and second segments being contiguous so that the primary probes abut one another in 5' phosphate-3'hydroxyl relationship, and so that a ligase can covalently fuse or ligate the two probes into a fused product. In addition, a third (secondary) probe can hybridize to a portion of the first probe and a fourth (secondary) probe can hybridize to a portion of the second probe in a similar abutting fashion. Of course, if the target is initially double stranded, the secondary probes also will hybridize to the target complement in the first instance. Once the ligated strand of primary probes is separated from the target strand, it will hybridize with the third and fourth probes which can be ligated to form a complementary, secondary ligated product. It is important to realize that the ligated products are functionally equivalent to either the target or its complement. By repeated cycles of hybridization and ligation, amplification of the target sequence is achieved. This technique is described more

completely in EP-A- 320 308 to K. Backman published June 16, 1989 and EP-A-439 182 to K. Backman *et al.*, published July 31, 1991, both of which are incorporated herein by reference.

For amplification of mRNAs, it is within the scope of the present invention to reverse transcribe mRNA into cDNA followed by polymerase chain reaction (RT-PCR); or, to use a single enzyme for both steps as described in U.S. Patent No. 5,322,770, or reverse transcribe mRNA into cDNA followed by asymmetric gap ligase chain reaction (RT-AGLCR) as described by R. L. Marshall, *et al.*, *PCR Methods and Applications* 4: 80-84 (1994).

Other known amplification methods which can be utilized herein include but are not limited to the so-called "NASBA" or "3SR" technique described in *PNAS USA* 87:1874-1878 (1990) and also described in *Nature* 350 (No. 6313):91-92 (1991); Q-beta amplification as described in published European Patent Application (EPA) No. 4544610; strand displacement amplification (as described in G. T. Walker *et al.*, *Clin. Chem.* 42:9-13 (1996)) and European Patent Application No. 684315; and target mediated amplification, as described by PCT Publication WO 9322461.

In one embodiment, the present invention generally comprises the steps of contacting a test sample suspected of containing a target polynucleotide sequence with amplification reaction reagents comprising an amplification primer, and a detection probe that can hybridize with an internal region of the amplicon sequences. Probes and primers employed according to the method herein provided are labeled with capture and detection labels wherein probes are labeled with one type of label and primers are labeled with the other type of label. Additionally, the primers and probes are selected such that the probe sequence has a lower melt temperature than the primer sequences. The amplification reagents, detection reagents and test sample are placed under amplification conditions whereby, in the presence of target sequence, copies of the target sequence (an amplicon) are produced. In the usual case, the amplicon is double stranded because primers are provided to amplify a target sequence and its complementary strand. The double stranded amplicon then is thermally denatured to produce single stranded amplicon members. Upon formation of the single stranded amplicon members, the mixture is cooled to allow the formation of complexes between the probes and single stranded amplicon members.

As the single stranded amplicon sequences and probe sequences are cooled, the probe sequences preferentially bind the single stranded amplicon members. This finding is counterintuitive given that the probe sequences are generally selected to be shorter than the primer sequences and therefore have a lower melt temperature than the primers. Accordingly, the melt temperature of the amplicon produced by the primers should also have a higher melt temperature than the probes. Thus, as the mixture is cooled, the re-formation of the double stranded amplicon is expected. As previously stated, however, this is not the case. Probes have been found to preferentially bind the single stranded amplicon members. Moreover, this

preference of probe/single stranded amplicon binding exists even when the primer sequences are added in excess of the probes.

After the probe/single stranded amplicon member hybrids are formed, they are detected. Standard heterogeneous assay formats are suitable for detecting the hybrids using the detection labels and capture labels present on the primers and probes. The hybrids can be bound to a solid phase reagent by virtue of the capture label and detected by virtue of the detection label. In cases where the detection label is directly detectable, the presence of the hybrids on the solid phase can be detected by causing the label to produce a detectable signal, if necessary, and detecting the signal. In cases where the label is not directly detectable, the captured hybrids can be contacted with a conjugate, which generally comprises a binding member attached to a directly detectable label. The conjugate becomes bound to the complexes and the conjugates presence on the complexes can be detected with the directly detectable label. Thus, the presence of the hybrids on the solid phase reagent can be determined. Those skilled in the art will recognize that wash steps may be employed to wash away unhybridized amplicon or probe as well as unbound conjugate.

A test sample is typically anything suspected of containing a target sequence. Test samples can be prepared using methodologies well known in the art such as by obtaining a specimen from an individual and, if necessary, disrupting any cells contained therein to release target nucleic acids. The target sequence is either double stranded or single stranded. In the case where PCR is employed in this method, the ends of the target sequences are usually known. In cases where LCR or a modification thereof is employed in the preferred method, the entire target sequence is usually known. Typically, the target sequence is a nucleic acid sequence such as, for example, RNA or DNA.

Generally, two primers which are complementary to a portion of a target strand and its complement are employed in PCR. For LCR, four probes, two of which are complementary to a target sequence and two of which are similarly complementary to the targets complement, are generally employed. While the amplification primers initiate amplification of the target sequence, in some cases, the detection (or hybridization) probe is not involved in amplification. Detection probes are generally nucleic acid sequences or uncharged nucleic acid analogs such as, for example, peptide nucleic acids which are disclosed in International Patent Application WO 92/20702; morpholino analogs which are described in U.S. Patents Nos 5,185,444, 5,034,506, and 5,142,047; and the like. Depending upon the type of label carried by the probe, the probe is employed to capture or detect the amplicon generated by the amplification reaction. The probe is not involved in amplification of the target sequence and therefore may have to be rendered "non-extendable" in that additional dNTPs cannot be added to the probe. In and of themselves analogs usually are non-extendable and nucleic acid probes can be rendered non-extendable by modifying the 3' end of the probe such that the hydroxyl group is no longer

capable of participating in elongation. For example, the 3' end of the probe can be functionalized with the capture or detection label to thereby consume or otherwise block the hydroxyl group. Alternatively, the 3' hydroxyl group simply can be cleaved, replaced or modified. U.S. Patent Application Serial No. 07/049,061 filed April 19, 1993 describes modifications which can be used to render a probe non-extendable.

While the length of the primers and probes can vary, the probe sequences are selected such that they have a lower melt temperature than the primer sequences. Hence, the primer sequences are generally longer than the probe sequences. Typically, the primer sequences are in the range of between 20 and 50 nucleotides long, more typically in the range of between 20 and 30 nucleotides long. The typical probe is in the range of between 10 and 25 nucleotides long.

Alternatively, a probe may be involved in the amplifying a target sequence, via a process known as "nested PCR". In nested PCR, the probe has characteristics, which are similar to those of the first and second primers normally used for amplification (such as length, melting temperature etc.), and as such, may itself serve as a primer in an amplification reaction. Generally in nested PCR, a first pair of primers (P_1 and P_2) are employed to form primary extension products. One of the primary primers (for example, P_1) may optionally be a capture primer (i.e. linked to a member of a first reactive pair), whereas the other primary primer (P_2) is not. A secondary extension product is then formed using a probe (P_1') and a probe (P_2') which may also have a capture type label (such as a member of a second reactive pair) or a detection label at its 5' end. The probes are complementary to and hybridize at a site on the template near or adjacent the site where the 3' termini of P_1 and P_2 would hybridize if still in solution. Alternatively, a secondary extension product can be formed using the P_1 primer with the probe (P_2') or the P_2 primer with the probe (P_1') sometimes referred to as "hemi-nested PCR". Thus, a labeled primer/probe set generates a secondary product which is shorter than the primary extension product. Furthermore, the secondary product may be detected either on the basis of its size or via its labeled ends (by detection methodologies well known to those of ordinary skill in the art). In this process, probe and primers are generally employed in equivalent concentrations.

Various methods for synthesizing primers and probes are well known in the art. Similarly, methods for attaching labels to primers or probes are also well known in the art. For example, it is a matter of routine to synthesize desired nucleic acid primers or probes using conventional nucleotide phosphoramidite chemistry and instruments available from Applied Biosystems, Inc., (Foster City, CA), Dupont (Wilmington, DE), or Milligen (Bedford, MA). Many methods have been described for labeling oligonucleotides such as

the primers or probes of the present invention. Enzo Biochemical (New York, NY) and CLONTECH (Palo Alto, CA) both have described and commercialized probe labeling techniques. For example, a primary amine can be attached to a 3' oligo terminus using 3'-Amine-ON CPG™ (CLONTECH). Similarly, a primary amine can be attached to a 5' oligo terminus using Aminomodifier II® (CLONTECH). The amines can be reacted to various haptens using conventional activation and linking chemistries. In addition, copending applications US. Serial Nos. 625,566, filed December 11, 1990 and 630,908, filed December 20, 1990, teach methods for labeling probes at their 5' and 3' termini, respectively. Publications WO92/10505, published 25 June 1992 and WO 92/11388 published 9 July 1992 teach methods for labeling probes at their 5' and 3' ends, respectively. According to one known method for labeling an oligonucleotide, a label-phosphoramidite reagent is prepared and used to add the label to the oligonucleotide during its synthesis. See, for example, N.T. Thuong *et al.*, Tet. Letters 29(46):5905-5908 (1988); or J. S. Cohen *et al.*, published U.S. Patent Application 07/246,688 (NTIS ORDER No. PAT-APPL-7-246,688) (1989). Preferably, probes are labeled at their 3' and 5' ends.

Capture labels are carried by the primers or probes and can be a specific binding member which forms a binding pair with the solid phase reagent's specific binding member. It will be understood, of course that the primer or probe itself may serve as the capture label. For example, in the case where a solid phase reagent's binding member is a nucleic acid sequence, it may be selected such that it binds a complementary portion of the primer or probe to thereby immobilize the primer or probe to the solid phase. In cases where the probe itself serves as the binding member, those skilled in the art will recognize that the probe will contain a sequence or "tail" that is not complementary to the single stranded amplicon members. In the case where the primer itself serves as the capture label, at least a portion of the primer will be free to hybridize with a nucleic acid on a solid phase because the probe is selected such that it is not fully complementary to the primer sequence.

Another method provided by the present invention comprises contacting a test sample with a plurality of polynucleotides wherein at least one polynucleotide is provided herein, hybridizing the test sample with the plurality of polynucleotides and detecting the hybridization complexes. The hybridization complexes are identified and quantified to compile a profile which is indicative of vanilloid receptor expression. Expressed RNA sequences may further be detected by reverse transcription and amplification of the DNA product by procedures well-known in the art, including polymerase chain reaction (PCR).

Drug Screening and Gene Therapy.

The present invention also encompasses the use of gene therapy methods for the introduction of antisense vanilloid receptor gene derived molecules such as polynucleotides or oligonucleotides of the present invention into patients with conditions associated with abnormal expression of polynucleotides related to vanilloid receptor. These molecules, including antisense RNA and DNA fragments and ribozymes, are designed to inhibit the translation of a vanilloid receptor mRNA, and may be used therapeutically in the treatment of conditions associated with altered or abnormal expression of a vanilloid receptor polynucleotide.

Alternatively, the oligonucleotides described above can be delivered to cells by procedures in the art such that the antisense RNA or DNA may be expressed in vivo to inhibit production of a vanilloid receptor polypeptide in the manner described above. Antisense constructs to vanilloid receptor polynucleotides, therefore, reverse the action of vanilloid receptor transcripts.

The present invention also provides a method of screening a plurality of compounds for specific binding to a human vanilloid receptor polypeptide, or any fragment thereof, to identify at least one compound which specifically binds a human vanilloid receptor polypeptide. Such a method comprises the steps of providing at least one compound; combining the vanilloid receptor polypeptide with each compound under suitable conditions for a time sufficient to allow binding; and detecting a vanilloid receptor polypeptide binding to each compound. Such a method permits the identification of vanilloid receptor binding compounds which modulate (i.e. repress or activate) the activity of a vanilloid receptor.

Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes for the polypeptide of the present invention, is used to design an antisense RNA oligonucleotide of from 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription, thereby preventing transcription and the production of vanilloid receptor derived polypeptide. For triple helix, see, for example, Lee *et al.*, Nucl. Acids Res. 6:3073 (1979); Cooney *et al.*, Science 241:456 (1988); and Dervan *et al.*, Science 251:1360 (1991) The antisense RNA oligonucleotide hybridizes to the mRNA in vivo and either blocks translation of an mRNA molecule into the vanilloid receptor polypeptide or alters the transport or stability of the mRNA. For antisense, see, for example, Okano, J. Neurochem. 56:560 (1991); and "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression", CRC Press, Boca Raton, Fla. (1988). Antisense oligonucleotides act with greater efficacy when modified to contain artificial internucleotide linkages which render the molecule resistant to nucleolytic cleavage. Such artificial internucleotide linkages include but are not limited to methylphosphonate, phosphorothiolate and phosphoroamydate internucleotide linkages.

The polypeptide or peptide fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids which can express the polypeptide or peptide fragment. Drugs may be screened against such transformed cells in competitive binding assays. For example, the formation of complexes between a polypeptide and the agent being tested can be measured in either viable or fixed cells.

The present invention thus provides methods of screening for drugs or any other agent which, upon binding to a vanilloid receptor, improves a condition or disease state. These methods comprise contacting the drug with a vanilloid receptor polypeptide or fragment thereof and assaying for the presence of a complex between the agent and the polypeptide. In competitive binding assays, a labeled agent, which is known to bind to the receptor, typically is used. After suitable incubation, free (or uncomplexed) agent is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the test agent or drug to bind to receptor polypeptide.

The present invention also encompasses the use of competitive drug screening assays in which neutralizing antibodies capable of binding receptor polypeptide specifically compete with a test drug for binding to the receptor polypeptide or fragment thereof. In this manner, the antibodies can be used to detect the presence of any polypeptide in the test sample which shares one or more antigenic determinants with a polypeptide provided herein.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to at least one receptor polypeptide disclosed herein. Briefly, large numbers of different small molecule test compounds or peptides are synthesized on a solid phase, such as plastic beads or some other surface. The test compounds are reacted with receptor polypeptide and washed. Test compounds thus bound to the solid phase are detected by methods well known in the art. Purified receptor polypeptide can also be coated directly onto plates for use in the drug screening techniques described herein. In addition, non-neutralizing antibodies can be used to capture the receptor polypeptide and immobilize it on the solid support. See, for example, EP 84/03564, published on September 13, 1984.

The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of the small molecules including agonists, antagonists, or inhibitors with which they interact. Such structural analogs can be used to fashion drugs which are more active or stable forms of the drug or which enhance or interfere with the function of the receptor polypeptide in vivo. (See, J. Hodgson, Bio/Technology 9:19-21 (1991).

For example, in one approach, the three-dimensional structure of a receptor polypeptide, or of a receptor polypeptide-inhibitor complex, is determined by x-ray crystallography, by

computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the receptor polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of a polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous polypeptide-like molecules or to identify efficient inhibitors.

Useful examples of rational drug design may include molecules which have improved activity or stability as shown by S. Braxton *et al.*, Biochemistry 31:7796-7801 (1992), or which act as inhibitors, agonists, or antagonists of native peptides as shown by S. B. P. Athauda *et al.*, J Biochem. (Tokyo) 113 (6):742-746 (1993).

It also is possible to isolate a target-specific antibody, selected by an assay as described hereinabove, and then to determine its crystal structure. In principle this approach yields a pharmacophore upon which subsequent drug design can be based. It further is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies ("anti-ids") to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-id is an analog of the original receptor. The anti-id then could be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides then can act as the pharmacophore (that is, a prototype pharmaceutical drug).

A sufficient amount of a recombinant receptor polypeptide of the present invention may be made available to perform analytical studies such as X-ray crystallography. In addition, knowledge of the polypeptide amino acid sequence which are derivable from the nucleic acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

Antibodies specific to a human vanilloid receptor polypeptide may further be used to inhibit its biological action by binding to the receptor polypeptide. In this manner, the antibodies may be used in therapy, for example, to treat disorders involving capsaicin-sensitive ion channels.

Further, such antibodies can detect the presence or absence of a human vanilloid receptor polypeptide and, therefore, are useful as diagnostic markers for the diagnosis of disorders involving vanilloid receptors. The present invention also is directed to antagonists and inhibitors of the receptor polypeptides of the present invention. The antagonists and inhibitors are those which inhibit or eliminate the function of the polypeptide. Thus, for example, an antagonist may bind to a receptor polypeptide of the present invention and inhibit or eliminate its function.

The antagonists and inhibitors may be employed as a composition with a pharmaceutically acceptable carrier, including but not limited to saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof.

Recombinant Technology.

The present invention provides vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the present invention and the production of polypeptides of the present invention by recombinant techniques. Such methods comprise culturing the host cells under conditions suitable for the expression of a human vanilloid receptor polynucleotide and recovering the polypeptide produced therefrom from the cell culture.

a. Host Cells

In one embodiment, the present invention provides host cells containing a recombinant construct as described below. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or a prokaryotic cell, such as a bacterial cell. Representative examples of appropriate hosts include bacterial cells, such as E. coli, Bacillus subtilis, Salmonella typhimurium; and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a routine matter of choice; fungal cells, such as yeast; insect cells such as Drosophila and Sf9; animal cells such as CHO, COS or Bowes melanoma; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings provided herein.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be a cloning vector or an expression vector. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying a vanilloid receptor gene. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

b. Vectors and Expression Systems

The present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. Such vectors include chromosomal, non-chromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, and viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. In a preferred embodiment, a construct comprises an expression vector (as described below). Large numbers of suitable plasmids and

vectors are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example: (a) Bacterial: pBR322 (ATCC 37017); pGEM1 (Promega Biotec, Madison, WI), pUC, pSPORT1 and pProEx1 (Life Technologies, Gaithersburg, MD); pQE70, pQE60, pQE-9 (Qiagen); pBs, phagescript, psiX174, pBluescript SK, pBsKS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5, and pGEX4T (Pharmacia Fine Chemicals, Uppsala, Sweden); and (b) Eukaryotic: pWLneo, pSV2cat, pOG44, pXT1, pSG (Stratagene); pSVK3, pBPV, pMSG, pSVL (Pharmacia); pcDNA3.1 (Invitrogen). Other appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, *supra*. Generally however, any plasmid or vector may be used as long as it is replicable and viable in a host.

In a preferred embodiment, the construct is an expression vector which also comprises regulatory sequences operably linked to the sequence of interest, to direct mRNA synthesis and polypeptide production. Regulatory sequences known to operate in prokaryotic and/or eukaryotic cells include inducible and non-inducible promoters for regulating mRNA transcription, ribosome binding sites for translation initiation, stop codons for translation termination and transcription terminators and/or polyadenylation signals. In addition, an expression vector may include appropriate sequences for amplifying expression (such as a dihydrofolate reductase gene).

Promoter regions may be selected from any desired gene but preferably from one which is highly expressed. Particular named bacterial promoters include lacZ, gpt, lambda P sub R, P sub L and trp. Eukaryotic promoters include cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, early and late SV40, LTRs from retroviruses, mouse metallothionein-I, prion protein and neuronal specific enolase (NSE). Selection of the appropriate promoter is well within the level of ordinary skill in the art. In addition, a recombinant expression vector will include an origin of replication and selectable marker (such as a gene conferring resistance to an antibiotic (e.g. neomycin, chloramphenicol or ampicillin) or a reporter gene (e.g. luciferase)) which permit selection of stably transformed or transfected host cells.

In a preferred prokaryotic or yeast expression vector, a heterologous structural sequence (i.e. a polynucleotide of the present invention) is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence will encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Preferred eukaryotic expression vectors will also comprise an origin of replication, a suitable promoter operably linked to a sequence of interest and also any necessary translation enhancing sequence, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Such vectors may also include an enhancer sequence to increase transcription of a gene. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription rate. Examples include the SV40 enhancer on the late side of the replication origin (bp 100 to 270), a cytomegalovirus early promoter enhancer, a polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

i. Vector construction

The appropriate DNA sequence may be inserted into a vector by a variety of procedures. Generally, site-specific DNA cleavage is performed by treating the DNA with suitable restriction enzymes under conditions which are generally specified by the manufacturer of these commercially available enzymes. Usually, about 1 microgram (μg) of plasmid or DNA sequence is cleaved by 1 unit of enzyme in about 20 microliters (μL) of buffer solution by incubation at 37°C for 1 to 2 hours. After incubation with the restriction enzyme, protein is removed by phenol/chloroform extraction and the DNA recovered by precipitation with ethanol. The cleaved fragments may be separated using polyacrylamide or agarose gel electrophoresis, according to methods known by the routine practitioner. (See Sambrook *et al.*, *supra*).

Ligations are performed using standard buffer and temperature conditions and with a ligase (such as T4 DNA ligase) and ATP. Sticky end ligations require less ATP and less ligase than blunt end ligations. When vector fragments are used as part of a ligation mixture, the vector fragment often is treated with bacterial alkaline phosphatase (BAP) or calf intestinal alkaline phosphatase (CIAP) to remove the 5'-phosphate and thus prevent religation of the vector. Alternatively, restriction enzyme digestion of unwanted fragments can be used to prevent ligation. Ligation mixtures are transformed into suitable cloning hosts such as E. coli and successful transformants selected by methods including antibiotic resistance, and then screened for the correct construct.

ii. Transformation/Transfection

Transformation or transfection of an appropriate host with a construct of the invention, such that the host produces recombinant polypeptides, may also be performed in a variety of ways. For example, a construct may be introduced into a host cell by calcium chloride transformation, lithium chloride or calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation. These and other methods for transforming/transfecting host

cells are well known to routine practitioners (see L. Davis *et al.*, "Basic Methods in Molecular Biology", 2nd edition, Appleton and Lang, Paramount Publishing, East Norwalk, CT (1994)).

iii. Recovery of Expressed Proteins from Recombinant Host Cells

Following transformation or transfection of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is derepressed by appropriate means (e.g., temperature shift or chemical induction), and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means (to release intracellular protein) and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents; such methods are well-known to the ordinary artisan. When the expressed protein has been secreted, it can be purified directly from the supernatant of harvested cells.

Vanilloid receptor polypeptide is recovered and purified from the supernatant or crude extract by known methods including ammonium sulfate or ethanol precipitation, acid extraction, affinity chromatography, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, hydroxyapatite chromatography or lectin chromatography. It is preferred to have low concentrations (approximately 0.1-5 mM) of calcium ion present during purification (Price, *et al.*, J. Biol. Chem. 244:917 (1969)). Protein refolding steps can be used, as necessary, in completing configuration of the protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

An alternative method for the production of large amounts of secreted protein involves the transformation of mammalian embryos and the recovery of the recombinant protein from milk produced by transgenic cows, goats, sheep, etc. Polypeptides and closely related molecules may be expressed recombinantly in such a way as to facilitate protein purification. One approach involves expression of a chimeric protein which includes one or more additional polypeptide domains not naturally present on human polypeptides. Such purification-facilitating domains include, but are not limited to, metal-chelating peptides such as histidine-tryptophan domains that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle, WA). The inclusion of a cleavable linker sequence such as Factor XA or enterokinase from Invitrogen (San Diego, CA) between the polypeptide sequence and the purification domain may be useful for recovering the polypeptide.

Immunoassays.

The polypeptides including their fragments or derivatives or analogs thereof of the present invention, or cells expressing them, can be used in a variety of assays, many of which are described herein, for the detection of antibodies to a human vanilloid receptor. They also can be used as an immunogen to produce antibodies. These antibodies can be, for example, polyclonal or monoclonal antibodies, chimeric, single chain and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

For example, antibodies generated against a polypeptide corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptide into an animal or by administering the polypeptide to an animal such as a mouse, rabbit, chicken, or goat. A mouse, rabbit or goat is preferred. The antibody so obtained then will bind the polypeptide itself. In this manner, even a sequence encoding only a fragment of the polypeptide can be used to generate antibodies that bind the native polypeptide. Such antibodies can then be used to isolate the polypeptide from test samples such as tissue suspected of containing that polypeptide. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique as described by Kohler and Milstein, *Nature* **256:** 495-497 (1975), the trioma technique, the human B-cell hybridoma technique as described by Kozbor *et al.*, *Immun. Today* **4:** 72 (1983), and the EBV-hybridoma technique to produce human monoclonal antibodies as described by Cole, *et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc, New York, NY, pp. 77-96 (1985). Techniques described for the production of single chain antibodies can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. See, for example, U.S. Pat. No. 4,946,778, which is incorporated herein by reference.

Various assay formats may utilize the antibodies of the present invention, including "sandwich" immunoassays and probe assays. For example, the monoclonal antibodies or fragment thereof of the present invention can be employed in various assay systems to determine the presence, if any, of a vanilloid receptor derived polypeptide in a test sample. For example, in a first assay format, a polyclonal or monoclonal antibody or fragment thereof, or a combination of these antibodies, which has been coated on a solid phase, is contacted with a test sample, to form a first mixture. This first mixture is incubated for a time and under conditions sufficient to form antigen/antibody complexes. Then, an indicator reagent comprising a monoclonal or a polyclonal antibody or a fragment thereof, or a combination of these antibodies, to which a signal generating compound has been attached, is contacted with the antigen/antibody complexes to form a second mixture. This second mixture then is

incubated for a time and under conditions sufficient to form antibody/antigen/antibody complexes. The presence of a vanilloid receptor derived polypeptide antigen present in the test sample and captured on the solid phase, if any, is determined by detecting the measurable signal generated by the signal generating compound. The amount of vanilloid receptor derived polypeptide antigen present in the test sample is proportional to the signal generated.

Or, a polyclonal or monoclonal vanilloid receptor derived polypeptide antibody or fragment thereof, or a combination of these antibodies which is bound to a solid support, the test sample and an indicator reagent comprising a monoclonal or polyclonal antibody or fragments thereof, which specifically binds to a vanilloid receptor derived polypeptide antigen, or a combination of these antibodies to which a signal generating compound is attached, are contacted to form a mixture. This mixture is incubated for a time and under conditions sufficient to form antibody/antigen/antibody complexes. The presence, if any, of a vanilloid receptor derived polypeptide present in the test sample and captured on the solid phase is determined by detecting the measurable signal generated by the signal generating compound. The amount of vanilloid receptor derived polypeptide proteins present in the test sample is proportional to the signal generated.

In another assay format, one or a combination of at least two monoclonal antibodies of the invention can be employed as a competitive probe for the detection of antibodies to a vanilloid receptor derived polypeptide protein. For example, vanilloid receptor derived polypeptide proteins such as the recombinant antigens disclosed herein, either alone or in combination, are coated on a solid phase. A test sample suspected of containing antibody to a vanilloid receptor derived polypeptide antigen then is incubated with an indicator reagent comprising a signal generating compound and at least one monoclonal antibody of the invention for a time and under conditions sufficient to form antigen/antibody complexes of either the test sample and indicator reagent bound to the solid phase or the indicator reagent bound to the solid phase. The reduction in binding of the monoclonal antibody to the solid phase can be quantitatively measured.

In yet another detection method, each of the monoclonal or polyclonal antibodies of the present invention can be employed in the detection of vanilloid receptor derived polypeptide antigens in fixed tissue sections, as well as fixed cells by immunohistochemical analysis. Cytochemical analysis wherein these antibodies are labeled directly (with, for example, fluorescein, colloidal gold, horseradish peroxidase, alkaline phosphatase, etc.) or are labeled by using secondary labeled anti-species antibodies (with various labels as exemplified herein) to track the histopathology of disease also are within the scope of the present invention.

In addition, these monoclonal antibodies can be bound to matrices similar to CNBr-activated Sepharose and used for the affinity purification of specific vanilloid receptor derived

polypeptide proteins from cell cultures or biological tissues such as to purify recombinant and native vanilloid receptor derived polypeptide antigens and proteins.

The monoclonal antibodies of the invention can also be used for the generation of chimeric antibodies for therapeutic use, or other similar applications.

The monoclonal antibodies or fragments thereof can be provided individually to detect vanilloid receptor derived polypeptide antigens. Combinations of the monoclonal antibodies (and fragments thereof) provided herein also may be used together as components in a mixture or "cocktail" of at least one vanilloid receptor derived polypeptide antibody of the invention with antibodies to other vanilloid receptor derived polypeptide regions, each having different binding specificities. Thus, this cocktail can include the monoclonal antibodies which are directed to different antigenic determinants of vanilloid receptor derived polypeptide proteins.

The polyclonal antibody or fragment thereof which can be used in the assay formats should specifically bind to a vanilloid receptor derived polypeptide region or other vanilloid receptor derived polypeptide protein used in the assay. The polyclonal antibody used is preferably of mammalian origin (such as from human, goat, rabbit or sheep). Most preferably, the polyclonal antibody is rabbit polyclonal anti-vanilloid receptor derived polypeptide antibody. The polyclonal antibodies used in the assays can be used either alone or as a cocktail of polyclonal antibodies.

It is contemplated and within the scope of the present invention that a vanilloid receptor derived polypeptide may be detectable in assays by use of a recombinant antigen as well as by use of a synthetic peptide or purified peptide, which contains an amino acid sequence of a vanilloid receptor derived polypeptide. It also is within the scope of the present invention that different synthetic, recombinant or purified peptides identifying different epitopes of a vanilloid receptor derived polypeptide can be used in combination in an assay to diagnose, evaluate, or prognose conditions associated with abnormal vanilloid receptor production. In this case, these peptides can be coated onto one solid phase, or each separate peptide may be coated on separate solid phases, such as microparticles, and then combined to form a mixture of peptides which can be later used in assays. Furthermore, it is contemplated that multiple peptides which define epitopes from different polypeptides may be used in combination to make a diagnosis, evaluation, or prognosis of abnormal vanilloid receptor production. To accomplish this, peptides coated on solid phases or labeled with detectable labels are allowed to compete with peptides from a patient sample for a limited amount of antibody. A reduction in binding of the synthetic, recombinant, or purified peptides to the antibody (or antibodies) is an indication of the presence of vanilloid receptor -secreted polypeptides in the patient sample where it may not be expected (for example, in cerebral spinal fluid). Such variations of assay formats are known to those of ordinary skill in the art.

In another assay format, the presence of antibody and/or antigen to vanilloid receptor derived polypeptide can be detected in a simultaneous assay, as follows. A test sample is simultaneously contacted with a capture reagent of a first analyte, wherein said capture reagent comprises a first binding member specific for a first analyte attached to a solid phase and a capture reagent for a second analyte, wherein said capture reagent comprises a first binding member for a second analyte attached to a second solid phase, to thereby form a mixture. This mixture is incubated for a time and under conditions sufficient to form capture reagent/first analyte and capture reagent/second analyte complexes. These so-formed complexes then are contacted with an indicator reagent comprising a member of a binding pair specific for the first analyte labeled with a signal generating compound and an indicator reagent comprising a member of a binding pair specific for the second analyte labeled with a signal generating compound to form a second mixture. This second mixture is incubated for a time and under conditions sufficient to form capture reagent/first analyte/indicator reagent complexes and capture reagent/second analyte/indicator reagent complexes. The presence of one or more analytes is determined by detecting a signal generated in connection with the complexes formed on either or both solid phases as an indication of the presence of one or more analytes in the test sample. In this assay format, recombinant antigens derived from human expression systems may be utilized as well as monoclonal antibodies produced from the proteins derived from the mammalian expression systems as disclosed herein. Such assay systems are described in greater detail in EP Publication No. 0473065.

In yet other assay formats, the polypeptides disclosed herein may be utilized to detect the presence of anti- vanilloid receptor derived polypeptide in test samples. For example, a test sample is incubated with a solid phase to which at least one recombinant protein has been attached. These are reacted for a time and under conditions sufficient to form antigen/antibody complexes. Following incubation, the antigen/antibody complex is detected. Indicator reagents may be used to facilitate detection, depending upon the assay system chosen. In another assay format, a test sample is contacted with a solid phase to which a recombinant protein produced as described herein is attached and also is contacted with a monoclonal or polyclonal antibody specific for the protein, which preferably has been labeled with an indicator reagent. After incubation for a time and under conditions sufficient for antibody/antigen complexes to form, the solid phase is separated from the free phase, and the label is detected in either the solid or free phase as an indication of the presence of a vanilloid receptor derived polypeptide antibody. Other assay formats utilizing the recombinant antigens disclosed herein are contemplated. These include contacting a test sample with a solid phase to which at least one antigen from a first source has been attached, incubating the solid phase and test sample for a time and under conditions sufficient to form antigen/antibody complexes, and then contacting the solid phase with a labeled antigen, which antigen is derived from a second source different from the first

source. For example, a recombinant protein derived from a first source such as E. coli is used as a capture antigen on a solid phase, a test sample is added to the so-prepared solid phase, and a recombinant protein derived from a different source (i.e., non-E. coli) is utilized as a part of an indicator reagent. Likewise, combinations of a recombinant antigen on a solid phase and synthetic peptide in the indicator phase also are possible. Any assay format which utilizes an antigen specific for a vanilloid receptor derived polypeptide from a first source as the capture antigen and an antigen specific for vanilloid receptor derived polypeptide from a different second source are contemplated. Thus, various combinations of recombinant antigens, as well as the use of synthetic peptides, purified proteins, and the like, are within the scope of this invention. Assays such as this and others are described in U.S. Patent No. 5,254,458, which enjoys common ownership and is incorporated herein by reference.

Other embodiments which utilize various other solid phases also are contemplated and are within the scope of this invention. For example, ion capture procedures for immobilizing an immobilizable reaction complex with a negatively charged polymer (described in EP publication 0326100 and EP publication No. 0406473), can be employed according to the present invention to effect a fast solution-phase immunochemical reaction. An immobilizable immune complex is separated from the rest of the reaction mixture by ionic interactions between the negatively charged poly-anion/immune complex and the previously treated, positively charged porous matrix and detected by using various signal generating systems previously described, including those described in chemiluminescent signal measurements as described in EPO Publication No. 0 273,115.

Also, the methods of the present invention can be adapted for use in systems which utilize microparticle technology including in automated and semi-automated systems wherein the solid phase comprises a microparticle (magnetic or non-magnetic). Such systems include those described in published EPO applications Nos. EP 0 425 633 and EP 0 424 634, respectively.

The use of scanning probe microscopy (SPM) for immunoassays also is a technology to which the monoclonal antibodies of the present invention are easily adaptable. In scanning probe microscopy, in particular in atomic force microscopy, the capture phase, for example, at least one of the monoclonal antibodies of the invention, is adhered to a solid phase and a scanning probe microscope is utilized to detect antigen/antibody complexes which may be present on the surface of the solid phase. The use of scanning tunneling microscopy eliminates the need for labels which normally must be utilized in many immunoassay systems to detect antigen/antibody complexes. The use of SPM to monitor specific binding reactions can occur in many ways. In one embodiment, one member of a specific binding partner (analyte specific substance which is the monoclonal antibody of the invention) is attached to a surface suitable for scanning. The attachment of the analyte specific substance may be by adsorption to a test

piece which comprises a solid phase of a plastic or metal surface, following methods known to those of ordinary skill in the art. Or, covalent attachment of a specific binding partner (analyte specific substance) to a test piece which test piece comprises a solid phase of derivatized plastic, metal, silicon, or glass may be utilized. Covalent attachment methods are known to those skilled in the art and include a variety of means to irreversibly link specific binding partners to the test piece. If the test piece is silicon or glass, the surface must be activated prior to attaching the specific binding partner. Also, polyelectrolyte interactions may be used to immobilize a specific binding partner on a surface of a test piece by using techniques and chemistries. The preferred method of attachment is by covalent means. Following attachment of a specific binding member, the surface may be further treated with materials such as serum, proteins, or other blocking agents to minimize non-specific binding. The surface also may be scanned either at the site of manufacture or point of use to verify its suitability for assay purposes. The scanning process is not anticipated to alter the specific binding properties of the test piece.

While the present invention discloses the preference for the use of solid phases, it is contemplated that the reagents such as antibodies, proteins and peptides of the present invention can be utilized in non-solid phase assay systems. These assay systems are known to those skilled in the art, and are considered to be within the scope of the present invention.

It is contemplated that the reagent employed for the assay can be provided in the form of a test kit with one or more containers such as vials or bottles, with each container containing a separate reagent such as a probe, primer, monoclonal antibody or a cocktail of monoclonal antibodies, or a polypeptide (either recombinant or synthetic) employed in the assay. Other components such as buffers, controls, and the like, known to those of ordinary skill in art, may be included in such test kits. It also is contemplated to provide test kits which have means for collecting test samples comprising accessible body fluids, e.g. blood, urine, saliva, and stool. Such collection means include lancets and absorbent paper or cloth for collecting and stabilizing blood; swabs for collecting and stabilizing saliva; cups for collecting and stabilizing urine or stool samples. Collection materials, papers, cloths, swabs, cups and the like, may optionally be treated to avoid denaturation or irreversible adsorption of the sample. The collection materials also may be treated with or contain preservatives, stabilizers or antimicrobial agents to help maintain the integrity of the specimens. Test kits designed for the collection, stabilization, and preservation of test specimens obtained by surgery or needle biopsy are also useful. It is contemplated that all kits may be configured in two components; one component for collection and transport of the specimen, and the other component for the analysis of the specimen. Further, kits for the collection, stabilization, and preservation of test specimens may be configured for use by untrained personnel and may be available in the open

market for use at home with subsequent transportation to a laboratory for analysis of the test sample.

The present invention will now be described by way of examples, which are meant to illustrate, but not to limit, the scope of the present invention. All references to the literature, including patents, patent applications and scientific publications, both hereinafter and above, are incorporated by reference in their entirety.

EXAMPLES

Example 1: Isolation of Full Length cDNA Clones of Human Vanilloid receptor (hVR1)

The amino acid sequence for the rat vanilloid receptor (Caterina *et al.*, (1997) *supra*) was used to search the LifeSeq™ human expression database (Incyte Pharmaceuticals, Inc., Palo Alto, CA) for human vanilloid receptor (hVR) sequences. A search performed with the rat sequence (using both BLAST and Smith Waterman algorithms) identified two overlapping ESTs (1427917 and 3460342) highly homologous to the rat sequence. These two ESTs overlapped to form a consensus sequence of 270 nt; 1427917 contained the DNA sequence from position 1-227 while 3460342 contained the DNA sequence from position 32-270. The consensus sequence derived from the two ESTs was compared with the published rat VR1 amino acid sequence (SEQ ID NO:4) using the GCG FRAMEALIGN program (FIG. 1) and showed 89 % identity.

PCR primers were prepared from the consensus sequence of ESTs 1427917 and 3460342 to clone the full length gene by RACE PCR (Frohman, MA (1991) Methods Enzymology 218:340-362). Two antisense RACE primers corresponding to nucleotide (nt) positions 59-86 and 120-140 from the aligned ESTs (1755 to 1782 and 1816-1836 of SEQ ID NO:7 respectively) were used to prime a human small intestine cDNA library and PCR products were isolated according to the manufacturer's instructions (Marathon Ready cDNA, CLONTECH). cDNAs were obtained which extended 1 kb upstream of the primers. An additional RACE primer was synthesized from the 5'- region of this PCR product (corresponding to nucleotides 855-883 of SEQ ID NO:7) and used to extend the cDNA clones upstream of the translation initiation codon. The sequences down stream of the EST consensus sequence were determined by DNA sequence analysis of the two Incyte clones. The final nucleic acid consensus sequence termed hVR1, (SEQ ID NO:7) and deduced amino acid sequence, (SEQ ID NO:8) were aligned using the GAP Program (Genetics Computer Group, Version 9, University of Wisconsin) with the rat gene and polypeptide sequences (SEQ ID NOs:2 and 4, respectively) and shown to have a DNA sequence identity of 82% (GAP Program, FIG. 2) and an amino acid sequence identity of 86% (GAP Program, FIG. 3). As shown in

FIG. 3, the amino acid sequence from positions 100-800 of SEQ ID NO:8 is 91 % identical to the rat gene product illustrating the sequence divergence at the amino terminal and carboxy terminal ends of the protein.

PCR primers were designed to amplify the hVR1 coding sequence while maintaining the relatively good Kozak consensus sequence for translation (an "A" at position -3 relative to the translation start site) and to provide flanking restriction enzyme sites for subsequent cloning into an expression vector. The sequence of the primers is shown below:

5' Mlu Cap Rec TTTAAACGCGTAGGATGAAGAAATGGAGC (SEQ ID NO:5)
Mlu1 start
3' Sal Cap Rec TATATTGTCGACGTCCTCACTTCTCCCCG (SEQ ID NO:6)
Sal1 stop

The hVR1 open reading frame was amplified by PCR using CLONTECH (Palo Alto, CA) human small intestine cDNA as a template and Pfu polymerase (Stratagene®, La Jolla, CA) to ensure a high fidelity product. The resulting PCR product was approximately 2500 bp in length. The PCR product was digested with Mlu1 and Sal1 and cloned into the pCIneo mammalian expression vector (Promega) at the Mlu1 and Sal1 sites. In addition, the PCR product from the human intestine cDNA was sequenced directly.

DNA sequencing of individual clones following transfection of the ligated PCR insert and the vector revealed three forms of the hVR1 gene product. Clone hVR1-1 was identical to SEQ ID NO:7 while hVR1-5 and hVR1-13 each contained two nucleotide substitutions which resulted in amino acid changes in the translation product. Clone hVR1-5 contained a both a C to G substitution at position 1144 (designated C1144G) and a C to T substitution at position 1605 (C1605T) in SEQ ID NO:7. Clone hVR1-13 contained both a G1325T substitution and an A1952G substitution. The direct sequencing of the PCR product consistently identified two peaks at positions 1144, 1605 and 1952 (see FIG. 4) suggesting that these variations represent polymorphic gene products. In contrast, only a single peak was found at positions 1325 suggesting that these substitutions were likely to be cloning artifacts. The artifact change at nt position 1325 for clone hVR1-13 was corrected by site specific mutagenesis using the QuickChange™ Site-Directed Mutagenesis Kit following the protocols of the manufacturer (Stratagene®) and resulted in clones hVR1-13.1.

The polymorphic changes also resulted in amino acid changes in the hVR1 protein. Specifically, the polymorphic variations at nucleotide positions 1144 and 1605 in clone hVR1-5 resulted two amino acid substitutions at position 315 (a methionine rather than an isoleucine) and 469 (an isoleucine rather than a threonine) of SEQ ID NO:8 whereas in clone hVR1-13.1,

the variation at nucleotide position 1952 resulted in a protein containing a valine in place of isoleucine at amino acid position 586 of SEQ ID NO:8.

Example 2: Additional Vanilloid Recepotor Homologs.

The Incyte and dbEST data bases were further searched using Blast and Smith-Waterman algorithms for homologies using the rat and human VR1 amino acid sequences. An additional Incyte EST (1682513) was found with homology in the proposed poor-loop region of rVR1 (Caterina et al., 1997, *supra*). Sequencing clone 1682513 and subsequent RACE cloning from kidney cDNA libraries identified a 3599 cDNA sequence (SEQ ID NO:1) containing an open reading frame (position 435-3050) coding for an 871 amino acid protein (SEQ ID NO:3) which we termed hVR3.

An alignment of hVR3 with hVR1, rVR1 and hVRL1 (Caterina et al, Nature 398:436-441 1999) using the GCG Pileup program with default parameters shows significant homology of hVR3 with the other ion channel members (FIG. 3). Particularly significant homology is found in the following structural domains all four sequences (i.e. hVR-1, rVR-1, hVR-2 and hVR-3):

a) three ankaryn repeat domains, at position 239-270, 294-317 and 370-403 in of FIG. 3

b) six hydrophobic regions consistent with transmembrane domains, positions 471-493, 519-540, 555-574, 579-597, 621-640 and 703-730

b) homology in the poor-loop region thought to mediate ion transport, position 671-691

The GAP analysis of hVR1 (bottom sequence) and hVR3 (top sequence) DNA sequences (SEQ ID NOs:7 and 1 respectively) showed 55% sequence identity (FIG. 5) while the GAP analysis of the derived amino acid sequences of hVR1 (bottom sequence) and hVR3 (top sequence), SEQ ID NOs:8 and 3 respectively, showed 46% sequence identity (FIG. 6).

Example 3: Quantitative RT-PCR.

Tissue distribution of hVR1 and hVR3 were determined by quantitative PCR (Q-PCR) using the ABI Prism 7700 following the recommendations of the manufacturer (PE-Applied Biosystems, Foster City, CA). The version of Q-PCR we utilized is referred to as TaqMan PCR.

In TaqMan PCR, a pair of amplification primers that hybridize to the target at sequence defined positions are utilized. In addition, a nucleic acid probe (labeled with a fluorophore and quencher molecule at either end, the TaqMan primer) is utilized that hybridizes to the target at a position between the two primers and preferably adjacent to one of the primers.

Importantly, this probe has a melting temperature that is significantly higher (8-10°C) than that of the primers. After the addition of the appropriate reagents (thermostable

polymerase, dNTPs, buffer, Mg2+) thermal cycling is begun and as in traditional PCR, the product doubles or nearly doubles each round in exponential fashion. Because the probe has a high melting temperature, during the course of the amplification reaction the polymerase encounters the probe bound tightly to the target at each hybridization/extension phase. The polymerase has a 5′-3′ nuclease activity that proceeds to degrade the obstruction and thereby release the fluorophore from the adjacent quencher molecule. After release, the fluorophore can be directly measured via laser excitation at the appropriate wave length. Consequently, the release of the probe fluorophore is directly linked to amplification and quantitative results can be generated by comparison to the amplification of appropriate standards of known concentration.

The amplification primers and nucleic acid probe used in Q-PCR were as follows:

Primer	Sequence	SEQ ID NO:
HVR1-Forward	5'-TGCCTGGAGCTGTTCAAGTTC-3'	SEQ ID NO:9
HVR1-Reverse	5'-TGATGAAGACAGCCTTGAAGTCA-3'	SEQ ID NO:10
HVR1-TaqMan	5'-AGTTCTCAGTGAACTCCAGGTCGCCCAT-3'	SEQ ID NO:11
HVR3-Forward	5'-ATTGAGAACCGCCACGAGAT-3'	SEQ ID NO:12
HVR3-Reverse	5'-AGACGGCCCCGAACTTG-3'	SEQ ID NO:13
HVR3-TaqMan	5'-CCATCAATGAACTGCTGCGGGACA-3'	SEQ ID NO:14

The hVR1 amplification primers corresponded to nucleotide positions 2102-2122 of SEQ ID NO:7 for the forward primer and nucleotide positions 2161-2183 of SEQ ID NO:7 for the reverse primer. The TaqMan primer corresponded to the reverse complement of nucleotide positions 2132-2159 of SEQ ID NO:7. The hVR3 forward, reverse and Taqman primers correspond to positions 1761-1780, 1826-1842 and 1795-1818 of SEQ ID NO:1 respectively. Ten μg total RNA was mixed with 5 μL of 50 ng/μL random hexamers in a final volume of 59 μL water, heated at 70° C for 10 min and placed on ice. The samples were adjusted to 20 mM Tris, pH 8.4, 50 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM dNTP and 800 units Superscript® II RT (Gibco BRL) in a 100 μL total volume and incubated at 25° C for 10 min., then 42° C for 50 min followed by 70° C for 10 min. and the samples were placed on ice. The samples were then incubated with 8 units of RNase H at 37° C for 20 min. The cDNA was purified and desalted by filtration on a Chromaspin column (CLONTECH) in TE and the samples quantified by OD_{260}. The Q-PCR reactions used 10 ng of cDNA template per 50 μL reaction containing 1X PCR bufferII with 600 nM ROX (PE-Applied Biosystems), 5 mM $MgCl_2$ and 1.25 units Amplitaq Gold (PE-Applied Biosystems). The samples were incubated at 95° C for 10 min. followed by 40 cycles of 95° C for 15 sec and 60° C for 1 min. Copy number was estimated by a dilution of plasmid DNA containing hVR1 or hVR3 and the cDNA

samples were normalized with Q-PCR reactions using 28S rRNA specific primers. The hVR1 and hVR3 RNA were expressed in relatively low amounts in all of the RNA samples tested (FIG. 7A and B). The hVR1 RNA was most abundant in brain, dorasal root ganglion (DRG), bladder, testes, and kidney while hVR3 was most abundant in kidney and bladder.

Example 4: Northern blotting assays.

Alternatively, or in addition to performing quantitative RT-PCR described in Example 2, the well known technique of Northern blotting provides for the detection of messenger RNA and gives a reasonable estimation of its size and steady-state level in a particular tissue (Sambrook *et al.*, *supra*). A Northern blotting assay may be performed as follows: Multiple Tissue Northern Blots are purchased from CLONTECH and probed with a vanilloid receptor cDNA fragment including all or part of SEQ ID NO:1. This fragment is labeled with α-^{32}P-dCTP by random priming using a commercial labeling kit (Stratagene®) to a specific activity of 1.1 x 10^9 cpm/ug DNA. The blots (membranes) are prehybridized at 60 °C for 1 hour in Express Hyb solution (supplied with the kit) and hybridized (also in Express Hyb solution) at the same temperature for two hours in the presence of denatured probe at 2 x 10^6 cpm/mL. After washing the blots twice in 2 x SSC + 0.5% SDS (20 min each wash), and twice under stringent conditions (0.1 x SSC + 0.01% SDS, 50 °C, 20 min. each wash), the filters are exposed to a phosphorimager screen.

Example 5: Ribonuclease Protection Assay

Alternatively, instead of or in addition to performing a Northern blot as described in Example 3, a ribonuclease protection assay may be performed as follows:

A. Labeling of Complementary RNA (cRNA) Hybridization Probes. Labeled sense and antisense riboprobes are transcribed from the EST sequence, which contains an RNA polymerase promoter such as SP6 or T7. The sequence may be from a vector containing the appropriate EST insert or from a PCR-generated product of the insert using PCR primers which incorporate an RNA polymerase promoter sequence. The transcripts are prepared in a 20 µL reaction volume containing 1 µg of DNA template, 2 µL of 100 mM dithiothreitol, 0.8 µL of RNasin (10-40U), 500 µM each of ATP, CTP, GTP, 5 µL (alpha^{32}P) UTP or 100-500 µM biotinylated UTP, and 1 µL of RNA polymerase in transcription buffer (40 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine HCl, 5 mM NaCl). Following incubation at 37°C for one hour, the transcripts are treated with DNase I (15 U) for an additional 30 min to digest the template. The probes then are isolated by spin columns, salt precipitation or electrophoresis

techniques which are well-known in the art. Finally, the probes are dissolved in lysis buffer (5 M guanidine thiocyanate, 0.1 M EDTA, pH 7.0).

B. Hybridization of Labeled Probe to Target. Approximately 20 μg of extracted total cellular RNA, prepared as described in Sambrook, *et al. supra*, is placed in 10 μL of lysis buffer and mixed with either (i) 1 x10^5 cpm of radioactively labeled probe or (ii) 250 pg of non-isotopically labeled probe, each in 2 μL of lysis buffer. The mixture then is incubated at 60°C for 5 min and hybridized overnight at room temperature. See, T. Kaabache *et al.*, *Anal. Biochem.* **232:** 225-230 (1995).

C. RNase Digestion. Hybridizations are terminated by incubation with 380 μL of a solution containing 40 μg/mL RNase A and 625 U/mL RNase T1 in 1 mM EDTA, 300 mM NaCl, 30 mM Tris-HCl pH 7.4 for 45-60 min at room temperature. RNase digestion then is terminated by the addition of 60 μL of proteinase-K (1.7 mg/mL) containing 3.3% SDS, followed by incubation for 30 min at 37°C. The digested mixture then is extracted with phenol:chloroform:isoamyl alcohol to remove protein. The mRNA:cRNA hybrids are precipitated from the aqueous phase by the addition 4 μg yeast tRNA and 800 μL of ethanol, and incubation at -80°C for 30 min. The precipitates are collected by centrifugation.

D. Fragment Analysis. The precipitates are dissolved in 5 μL of denaturing gel loading dye (80% formamide, 10 mM EDTA, pH 8.0, 1 mg/mL xylene cyanol, 1 mg/mL bromophenol blue) and electrophoresed in 6 % polyacrylamide TBE, 8 M urea denaturing gels. The gels are dried under vacuum and autoradiographed. Quantification can be performed by comparing the counts obtained from the test samples to a calibration curve that was generated by utilizing calibrators that are the sense strand. In cases where non-isotopic labels are used, hybrids are transferred from the gels to membranes (nylon or nitrocellulose) by blotting and then analyzed using detection systems that employ streptavidin alkaline phosphatase conjugates and chemiluminesence or chemifluoresence reagents. Again, expression of an mRNA which is detectable by the labeled probe in a particular tissue suggests that vanilloid receptor is produced in that tissue.

Example 6: Identification of Additional Members of the Vanilloid Receptor Family.

The Northern blot method described in Example 3 *supra* can detect distinct messages only if they have large differences in sizes (more than 100 to 200 nucleotides); small differences in message size (such as those arising from alternative splicing in the coding region) are not detected by this method. Instead, other strategies are used to detect possible variants of vanilloid receptor message and determine their steady-state levels. Splice variants in the coding region can be detected by RT-PCR using primers designed to give products of small size. Variants in the 3' UTR can also be detected by RT-PCR. In RT-PCR, the forward primer is

chosen in a region of the ORF that is common to all message variants known so far, as close as possible to the stop codon. The reverse primer is an oligo-dT anchored with a dinucleotide for the specificity. Since the first nucleotide of the anchor can be A, C, or G, and the second nucleotide can be either A, C, G, or T, a combination of 12 anchored reverse primers are needed. Each reverse primer is thus used with the unique forward primer, in 12 different reactions. The PCR products are then run in an agarose gel and detected by UV fluorescence after ethidium bromide staining. Because of its high sensitivity and specificity, this method allows the detection of even small size and sequence variations in the 3' UTR. Sequence variations at the 5' end of the mRNA can be found using the RACE-PCR technique with similar sensitivity of detecting variant products.

Example 7: Dot Blot/Slot Blot

Dot and slot blot assays are quick methods to evaluate the presence of a specific nucleic acid sequence in a complex mix of nucleic acid.

Up to 20 µg of RNA is mixed in 50 µL of 50% formamide, 7% formaldehyde, 1X SSC, and allowed to incubate 15 min at 68°C and cooled on ice. Then, 100 µL of 20X SSC is added to the RNA mixture and loaded onto a vacuum-manifold apparatus that has a prepared nitrocellulose or nylon membrane. The membrane is soaked in water, 20X SSC for 1 hour, placed on two sheets of 20X SSC prewet Whatman #3 filter paper, and inserted into a slot blot or dot blot vacuum manifold apparatus. The slot blot is analyzed with probes prepared and labeled as in Example 4 *supra*.

Other methods and buffers not specifically detailed for Examples 3-5 are described in Sambrook *et al.*, *supra*.

Example 8: In Situ Hybridization

This method is useful to directly detect specific target nucleic acid sequences in cells using detectable nucleic acid hybridization probes.

Tissues are prepared with cross-linking fixatives agents such as paraformaldehyde or glutaraldehyde for maximum cellular RNA retention. See, L. Angerer *et al.*, *Methods in Cell Biol.* **35:** 37-71 (1991). Briefly, the tissue is placed in greater than 5 volumes of 1% glutaraldehyde in 50 mM sodium phosphate, pH 7.5 at 4°C for 30 min. The solution is changed with fresh solution for a further 30 min fixing. The fixing solution should have an osmolality of approximately 0.375% NaCl. The tissue is washed once in isotonic NaCl to remove the phosphate.

The fixed tissues then are embedded in paraffin, as follows. The tissue is dehydrated through a series of ethanol concentrations for 15 min each: 50% twice, 70% twice, 85%, 90% and 100% twice. The tissue next is soaked in two changes of xylene for 20 min each at room

temperature; then it is soaked in two changes of 1 xylene:1 paraffin for 20 min each at 60°C; and then it is soaked in three final changes in paraffin for 15 min each.

The tissue next is cut in 5 μm sections using a standard microtome and placed on a slide previously treated with the tissue adhesive 3-aminopropyltriethoxysilane.

Paraffin is removed from the tissue by two 10 min xylene soaks and rehydrated in a series of ethanol concentrations; 99% twice, 95%, 85%, 70%, 50%, 30% and distilled water twice. The sections are pre-treated with 0.2 M HCl for 10 min and permeabilized with 2 μg/mL Proteinase-K at 37°C for 15 min.

Labeled riboprobes transcribed from the pSPORT1 plasmid containing fragments of vanilloid receptor cDNA are hybridized to the prepared tissue sections and hybridized overnight at 56°C in 3X standard saline extract and 50% formamide. Excess probe is removed by washing in 2X standard saline citrate and 50% formamide followed by digestion with 100 μg/mL RNase A at 37°C for 30 min. Fluorescence probe is visualized by illumination with UV light under a microscope. Fluorescence in the cytoplasm is indicative of mRNA production. Fluorescence in the nucleus detects the presence of genomic material. Alternatively, the sections can be visualized by autoradiography.

Example 9: Bacterial Expression and Purification of Human Vanilloid Receptor

A. Construction of Expression Vectors containing DNA Fragments Encoding hVR3: DNA fragments encoding hVR3 (containing all or part of SEQ ID NO:1) are generated by PCR for introduction into a prokaryotic expression vector such as pProExI, (Life Technologies, Gaithersburg, MD) using hVR-3 as template DNA. The primers are designed to allow the inframe insertion of the vanilloid receptor fragment with the prokaryotic translation initiation and His tagged regions. After amplification, the PCR products are digested with appropriate restriction enzymes, and ligated into pProExI, (previously digested with the same restriction enzymes) using standard ligation techniques (see J. Sambrook, *et al. supra*). *E. coli* DH5∝ cells are then transformed with the ligation mixtures and selected on medium containing ampicillin. Plasmid DNAs are prepared from individual clones and subjected to restriction enzyme analysis to confirm that the hVR3 inserts are in the proper orientation.

B. Purification of His-tagged hVR3: In the pProExI expression system, a desired protein is produced with a tag of six histidine residues fused upstream of the protein. Accordingly, the pProExI vectors containing the cloned hVR3 genes or gene fragments thereof are expected to produce fusion proteins of his-tagged hVR3 which could be purified by affinity chromatography to a nickel-conjugated resin. To produce the fusion proteins for purification, recombinant bacteria are grown overnight in Luria broth containing 50 μg/mL ampicillin (LB + amp) on a rotary shaker at 225 rpm, at 37°C and used to inoculate fresh LB + amp (300 mL) at

a 1:10 dilution. The fresh cultures are incubated, with shaking at 225 rpm, at 37°C for 1 hour, induced with isopropyl ß-D-thiogalactopyranoside (IPTG, 1 mM) and re-incubated for an additional 3 hours. Cultures are then centrifuged at 5,000 g to pellet the bacteria. Pellets are resuspended in 10 mL of lysis buffer (50 mM sodium phosphate (pH 8.0), 0.3 M NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 0.2 mM benzamidine) containing 1% TRITON-X100 at 4°C and sonicated on ice until greater than 90% of the cells are lysed (as determined by OD_{590}). After sonication, cell debris and unlysed cells are removed by centrifugation at 10,000 g for 10 minutes at 4°C. The resulting supernatant is loaded onto a 3 mL bed volume nickle-nitro-triacetic acid column (Ni-NTA,.QIAGEN, Chatsworth, CA) pre-equilibrated with 10 bed volumes of lysis buffer containing 0.1% TRITON-X100. The column is sequentially washed with 10 bed volumes of wash buffer (50 mM sodium phosphate, pH 6.0, 0.3 M NaCl, and 0.1% TRITON-X100) and 10 bed volumes of 50 mM imidazole in the same buffer (to remove non-specifically bound proteins). The his-tagged vanilloid receptor fusion proteins are eluted from the column with a total of 5 bed volumes of wash buffer containing 0.2 M imidazole and collected as 2 mL fractions. The purity of each eluted fusion protein is assessed after SDS-PAGE on a 13.5% gel stained with Coomassie blue. The protein concentration is determined by absorbance at 280nm after determining the relative extinction coefficients for each of the recombinant fusion proteins or by using the bicinchoninic acid procedure (BCA Protein Assay, Pierce).

Example 10: Expression of hVR3 in Eukaryotic Cells

Expression of hVR3 in mammalian cells was achieved by lipofection using the Lipofectamine Plus Reagent (Gibco BRL). HEK 293 cells (5×10^6) were plated into 10 cm culture dishes (Falcon 1005) in antibiotic free media (Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 2 mM Glutamine and 10% fetal bovine serum (FBS), Gibco BRL, Gaithersburg, MD) in a humidified atmosphere containing 5% CO_2 at 37°C and grown overnight to approximately 70-80% confluence. Following media replacement (8 mL of the antibiotic free medium), the cells were incubated with a mixture of DNA and Lipofectamine PLUS Reagent (17 µg DNA, 147 µL PLUS Reagent, 42 µL lipofectamine) diluted into 2.1 mLs of serum-free media and incubated for three-four hours. An additional 8 mLs of antibiotic free medium was then added and the cells were incubated overnight. For transient assays, the cells were replated into 96 well plates at 1×10^5 cells /well and channel activity was measured the following day (see example 10). For stable cell lines, the neomycin-resistant cells were selected with 800 µg/mL Geneticin (Gibco BRL) and the media was replaced frequently to remove dead cells and debris. Individual colonies were obtained by clonal selection, a technique well known in the art.

Example 11: Functional Analysis of hVR3 in Eukaryotic Cells

A. Functional characterization by FLIPR analysis: A fluorescent imaging plate reader (FLIPR, Molecular Devices, Sunnyvale CA) is used to assay changes in intracellular calcium in cells pre-loaded with the calcium sensitive dye Fluo-3 AM (Molecular Probes, Eugene, Oregon). Either transiently transfected HEK 293 cells or stably transfected subclones of these cells are grown to confluence in black-walled 96-well tissue culture plates in complete medium (antibiotic free medium described in Example 9 + 1% antibiotic/antimycotic (Gibco BRL)+ 200 μg/mL Geneticin). A Fluo-3 AM solution is prepared by dispersing 40 μL of a 1 mg/mL stock Fluo-3/DMSO solution into 10 mL Dulbecco's phosphate-buffered saline (D-PBS). The growth medium is replaced with the Fluo-3 solution and the cells are incubated in the dark at room temperature for 1 hr. The cells are washed gently 3 times with D-PBS using a Denley Cellwash instrument resulting in a final volume of 100 μL D-PBS per well.

Three different plates are loaded into the FLIPR for the assay: (1) the washed HEK 293 cell plate; (2) a plate containing D-PBS or vanilloid receptor antagonist at 4 times the desired concentration; and (3) a plate containing D-PBS or vanilloid receptor agonists at 4 times the desired concentration. The cells are assayed in FLIPR as follows. All pipetting steps are performed by the FLIPR's built-in pipetting armature: 50 μL from the antagonist/D-PBS plate is added to the cell plate 10 seconds after the start of the analysis and incubated for 5 min followed by the addition of 50 μL from the agonist plate and incubation for an additional 10 min. The FLIPR instrument collected fluorescence data throughout the course of the analysis.

B. Functional characterization using Xenopus oocytes expression: Xenopus oocytes are used for expression of hVR3 and measuring electrophysiological responses essentially as described by Briggs et al. (Neuropharmacology. 1995 Jun;34(6):583-590 [1995]). Female *Xenopus lævis* frogs are obtained from Nasco (Fort Atkinson, WI) and are maintained and treated using standard protocols approved by Abbott's Institutional Animal Care and Use Committee. Frogs are anesthetized with tricaine (0.28%) and sacrificed by decapitation and pithing. Ovaries are removed and placed in low-Ca^{2+} Barth's solution (87.5 mM NaCl, 2.5 mM KCl, 1 mM $MgCl_2$, and 10 mM Na-HEPES buffer, final pH 7.55). Sections of the ovaries can be maintained at 4°C for up to three weeks for additional oocyte preparations.

For each oocyte isolation, several ovary lobes are taken, opened using blunt dissection, rinsed in low-Ca^{2+} Barth's solution, and incubated in collagenase (Type 1A, Sigma Chemical Co., St. Louis, MO; 2 mg/mL in low-Ca^{2+} Barth's solution) for 1-2 hours at room temperature. Defolliculation is completed manually. The isolated oocytes are maintained at 17-18°C in normal Barth's solution (90 mM NaCl, 1 mM KCl, 0.66 mM $NaNO_3$, 2.4 mM $NaHCO_3$, 0.74 mM $CaCl_2$, 0.82 mM $MgCl_2$, 2.5 mM sodium-pyruvate, 10 mM Na-HEPES buffer (final pH 7.55), 100 U/ml penicillin and 100 μg/ml streptomycin). Oocytes are injected with 50 ng (50

nL of 1 μg/uL) hVR3 RNA within 24 hours of their preparation, and are used within 2-7 days after injection.

Recordings are made using two-electrode voltage clamp in a virtual-ground configuration (Geneclamp 500 amplifier, Axon Instruments, Foster City, CA). Electrodes are made from borosilicate glass (1.5 mM o.d., 1.17 mm i.d.) and are filled with 120 mM KCl. The impedance of the current-passing electrode is determined. Voltage as well as current are recorded to monitor the quality of the voltage clamp and voltage losses are kept below 2 mV.

Experiments are performed in OR2 solution containing 90 mM NaCl, 2.5 mM KCl, 2.5 mM $CaCl_2$, 1 mM $MgCl_2$, and 5 mM Na-HEPES buffer, final pH 7.4. The recording chamber (RC5/18, Warner Instruments, Hamden, CT) is perfused at 3 mL/min. Capsaicin is applied for a precisely controlled time using a push/pull applicator positioned to within 300-400 um from the oocyte and a solenoid valve controlled by the data acquisition system (Digidata 1200 A/D board and pClamp 6 software; Axon Instruments). Antagonists (capsazepine) are superfused in the bathing solution for > 3 minutes prior to testing agonist (capsaicin) in the presence of antagonist. Both antagonist and agonist are present in the ligand applicator so that the concentration of antagonist remains constant during agonist application.

C. Functional characterization using whole cell patch clamping of transfected mammalian cells: HEK-293 cells stably transfected with hVR3 are maintained in the same culture medium as used for FLIPR measurements, but are plated onto glass coverslips in 24-well culture dishes and are used at low cell-density so that recordings so that recordings can be made from individual cells. Whole-cell patch-clamp recordings are made using standard techniques (Axopatch 200B amplifier, Digidata 1200 A/D board and pClamp 6 software; Axon Instruments). Electrodes are made from Corning 7052 glass (1.65 mm o.d., 1.1 mm i.d.; Warner Instrument Corp., Hamden, CT) and have resistances of 1-4 MΩ. The external (bath) solution contains 145 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM dextrose, and 10 mM Na-HEPES buffer (final pH 7.4, 310 mOsm). The internal (pipette) solution contains 130 mM K-aspartate, 10 mM KCl, 10 mM K-BAPTA (Ca^{2+} chelator), 2 mM Mg-ATP, and 10 mM K-HEPES buffer (final pH 7.3, 280 mOsm).

Agonists and antagonists are applied to the recorded cell using a DAD-12 computer-controlled micro-superfusion system (ALA Scientific, Westbury, NY). Responses typically are recorded at a holding potential of -60 mV and are normalized to the response to 1 uM capsaicin as a standard to correct for variance of receptor expression from cell to cell. A protocol is devised wherein a compound can be tested in the midst of the standard 1 uM capsaicin application. After applying capsaicin for 20 seconds to determine the standard response, the micro-superfusion is switched immediately to the test compound (e.g., different concentration of capsaicin, or different agonist, or antagonist in the presence of capsaicin) for 20 seconds,

followed immediately by 1 uM capsaicin again for 20 seconds to test reversibility of the test compound, and finally back to normal bath solution.

Example 12: Production of Synthetic Peptides of human vanilloid receptor

Synthetic peptide sequences (15-25 amino acids) are selected from the hVR3 sequence (SEQ ID NO:8). Peptides are synthesized on an ABI Peptide Synthesizer (available from Applied Biosystems, Foster City, CA), Model 431A, using standard reagents and conditions known in the art for solid phase peptide synthesis (see for example, Stewart, J.M., and Young, D.J., *Solid Phase Peptide Synthesis*, W.H. Freeman Co., San Francisco, 1963). Briefly, a peptide sequence is generated on a resin (such as chloromethyl-polystyrene-divinylbenzene) by the sequential coupling of one or more amino acids or suitably protected amino acids to a growing peptide chain. Cleavage of the peptide from the resin and final deprotection of the peptide are achieved by adding the resin to 20 mL trifluoroacetic acid (TFA), 0.3 mL water, 0.2 mL ethanedithiol, 0.2 mL thioanisole and 100 mg phenol, and stirring at room temperature for 1.5 hours. The resin then is filtered by suction and the peptide obtained by precipitation of the TFA solution with ether, followed by filtration. Each peptide is purified via reverse-phase preparative HPLC using a water/acetonitrile/0.1% TFA gradient and lyophilized. The product is confirmed by mass spectrometry.

Example 13: Production of Polyclonal Antibodies to human vanilloid receptor.

A. Preparation of Immunizing Antigens: Purified synthetic peptides are prepared as described in Example 10. To generate antigens for immunization, the purified peptides are conjugated to Keyhole Limpet Hemocyanin (KLH) and bovine serum albumin (BSA) using an Imject Activated Immunogen Conjugation Kit (Pierce, Rockford, Il) in accordance with the manufacturer's instructions.

B. Immunization Protocol: Polyclonal antisera are generated using the protocol of the Berkeley Antibody Company (Berkeley, CA). Before receiving the first immunization, a sample of preimmune blood (5 mL) is drawn from each of at least 2 rabbits. Afterward, each rabbit is injected subcutaneously with an aliquot of KLH-conjugated peptide (200-500 μg) in Complete Freunds Adjuvant. After 21 days, the immune response is boosted with a second subcutaneous injection of KLH-conjugated peptide (100-250 μg) in Incomplete Freund's Adjuvant. Blood (50 mL) is collected on day 31 and serum tested for reactivity to BSA-coupled peptide using an enzyme linked immunoadsorbant assay (ELISA). Subsequent boosts with KLH-conjugated peptide are given on days 42, 63 and 84 (post injection #1) and production bleeds (50 mL) drawn on days 52, 73 and 94 for testing by ELISA in the manner described. Serum is then stored at -20°C until further use.

Example 14: Inhibition of endogenous vanilloid receptor expression.

Antisense RNA or DNA is a strategy currently widely used to reduce or completely block the endogenous synthesis of proteins . The antisense molecule can be an oligonucleotide targeted to a particular region of the endogenous message, or can be transcribed from an expression vector in which the cDNA for the target gene is ligated in the antisense orientation. In the case of vanilloid receptor, multiple antisense molecules (20-30 nt) are made spanning the complete mRNA and inhibition is measured by the reduction in steady state vanilloid receptor RNA levels in transfected cells using quantitative RT-PCR (Example 2). The oligonucleotides are ranked by their ability to inhibit expression and the best are used in further experiments.

We claim:

1. An isolated or purified polynucleotide derived from human vanilloid receptor, wherein said polynucleotide has at least 83% identity with SEQ ID NO:1, and complements thereof.

2. The polynucleotide of Claim 1 having the nucleotide sequence of SEQ ID NO:1.

3. The polynucleotide of Claim 1 wherein said polynucleotide has a sequence of from about nucleotide position 435 to about nucleotide position 3050 of SEQ ID NO:1.

4. The polynucleotide of Claim 1 wherein said nucleotide sequence selectively hybridizes to SEQ ID NO:1.

5. The polynucleotide of Claim 1 wherein said polynucleotide has an overall length of about 20 to about 50 nucleotides.

6. The polynucleotide of Claim 1 wherein said polynucleotide has an overall length of about 20 to about 25 nucleotides.

7. The polynucleotide of Claim 1, wherein said polynucleotide is produced by recombinant techniques.

8. The polynucleotide of Claim 1, wherein said polynucleotide is produced by synthetic techniques.

9. The polynucleotide of Claim 1 which encodes SEQ ID NO:3.

10. The polynucleotide of Claim 1, wherein said polynucleotide comprises a sequence encoding at least one human vanilloid receptor epitope.

11. The polynucleotide of Claim 1, wherein said polynucleotide is attached to a solid phase.

12. The polynucleotide of Claim 10, wherein said solid phase comprises an array of polynucleotide molecules attached thereto.

13. A recombinant expression system comprising a nucleic acid sequence that includes an open reading frame derived from a human vanilloid receptor polynucleotide, wherein said open reading frame is operably linked to a control sequence compatible with a desired host, and said nucleic acid sequence has at least 83% identity to SEQ ID NO:1, and fragments and complements thereof.

14. The recombinant expression vector of Claim 13 selected from the group consisting of pProEx1 and pCIneo.

15. A host cell transformed with the recombinant expression system of Claim 13.

16. The host cell of Claim 15 wherein said host cell is a eukaryotic cell.

17. A human vanilloid receptor polypeptide encoded by a nucleic acid sequence having at least 83% identity to SEQ ID NO:1.

18. The polypeptide of Claim 17 having SEQ ID NO:3.

19. The polypeptide of Claim 17 produced by recombinant techniques.

20. The polypeptide of Claim 17 produced by synthetic techniques.

21. A method for producing a polypeptide containing at least one human vanilloid receptor epitope comprising incubating host cells transformed with an expression vector wherein said expression vector comprises a nucleotide sequence having at least 83% identity to SEQ ID NO:1.

22. The method of Claim 21 wherein said nucleotide sequence which encodes a human vanilloid receptor has the sequence of SEQ ID NO:1.

23. The method of Claim 21 wherein said nucleotide sequence which encodes a human vanilloid receptor has the sequence of SEQ ID NO:1 from about nucleotide position 435 to about nucleotide position 3050.

24. The method of Claim 21 wherein said nucleotide sequence encodes a human vanilloid receptor having the sequence SEQ ID NO:3.

25. A method for identifying compounds that modulate vanilloid receptor activity, comprising:

(a) providing a host cell that expresses said vanilloid receptor polypeptide;

(b) mixing a test compound with said cell; and

(c) measuring either

(i) the effect of the test compound on the cell expressing the receptor, or

(ii) the binding of the test compound to the cell or to the receptor.

26. The method of claim 25, wherein said host cell is a eukaroyotic or procaryotic cell.

27. The method of claim 25, wherein said measurement of step (c) (ii) is performed by measuring a signal generated by a signal-generating compound.

28. The method of claim 25, wherein the measurement of step (c) (ii) is performed by measuring a signal generated by a radiolabeled ion, a fluorophore or an electrical current.

29. The method of claim 25, further comprising the step of pre-loading the host cell with a fluorescent dye prior to step (b).

30. The method of claim 29, wherein the measurement of step (c) (i) is performed by measuring a signal generated by the fluorescent dye.

31. A method for identifying a cytoprotective compound, comprising:

(a) providing a cell that expresses a vanilloid receptor polypeptide or fragment thereof;

(b) combining a test compound with the cell; and

(c) monitoring the cell or cellular function for an indication of cytotoxicity.

32. The method of Claim 31 wherein said cell is either a procaryotic or eukaroyotic cell.

33. The method of Claim 31, wherein said cell comprises an expression vector comprising the polynucleotide having at least 83% to SEQ ID NO:1 which is operably linked to control sequences that direct the transcription of the polynucleotide whereby the polynucleotide is expressed in a host cell.

34. The method of claim 31, wherein at least one of the control sequences comprises an inducible promotor.

35. The method of claim 34, wherein said cell is maintained in the presence of a substance such as to minimize or block a cytotoxic effect on said cell.

36. A method of treating an individual having a condition associated with vanilloid receptor modulation, comprising administering to said individual an effective amount of a compound that controls the gene expression of vanilloid receptor, in a pharmaceutically acceptable excipient.

37. A monoclonal antibody which specifically binds to human vanilloid receptor having amino acid sequence SEQ ID NO:3 or fragments thereof.

38. A polyclonal antibody which specifically binds to human vanilloid receptor having amino acid sequence SEQ ID NO:3 or fragments thereof.

```
  1 CGGCCGTCGATGAAGACCCTGTTTGTGGACAGCTAcAGTGAGATGCTTTTCTTTCTGCAG 60
    |||||||||:::|||...|||||||||||||||||||||||...|||||||||...|||
500 ArgProSerLeuLysSerLeuPheValAspSerTyrSerGluIleLeuPhePheValGln 519

 61 TCACTGTTCATGCTGGCCACCGTGGTGCTGTACTTCAGCCACCTCAAGGAGTATGTGGCT 120
    ||||||||||||||   ...||||||||||||||      ||||||||||||||
520 SerLeuPheMetLeuValSerValValLeuTyrPheSerGlnArgLysGluTyrValAla 539

121 TCCATGGTATTCTCCCTGGCCTTGGGCTGGACCAACATGCTCTACTACACCCGCGGTTTC 180
    |||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||
540 SerMetValPheSerLeuAlaMetGlyTrpThrAsnMetLeuTyrTyrThrArgGlyPhe 559

181 CAGCAGATGGGCATCTATGCCGTCATGATAGAGAAGATGATCCTGAGAGACCTGTGCCGT 240
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
560 GlnGlnMetGlyIleTyrAlaValMetIleGluLysMetIleLeuArgAspLeuCysArg 579

241 TTCATGTTTGTCTACATCGTCTTCTTGTTC 270
    ||||||||||||||:::||||||||||||
580 PheMetPheValTyrLeuValPheLeuPhe 589
```

FIG. 1

```
163 GCAGGTTGCACACTGGGCCACAGAGGATCCAGCAAGGATGAAGAAATGGA 212
    || |||||||| | ||||||||||||||||||| | |||||||| | ||  ||
 44 GCTGGTTGCAAATTGGGCCACAGAGGATCTGGAAAGGATGGAACAACGGG  93

213 GCAGCACAGACTTGGGGGCAGCTGCGGACCCACTCCAAAAGGACACCTGC 262
    |||  |||||| | ||   ||| | |||||| |||| || || ||||||
 94 CTAGCTTAGACTCAGAGGAGTCTGAGTCCCCACCCCAAGAGAACTCCTGC 143

263 CCAGACCCCCTGGATGGAGACCCTAACTCCAGGCCACCTCCAGCCAAGCC 312
    |  |||||  |  ||  |||||||||||||  || ||||||||||||  |||||||
144 CTGGACCCTCCAGACAGAGACCCTAACTGCAAGCCACCTCCAGTCAAGCC 193

313 CCAGCTCTCCACGGCCAAGAGCCGCACCCGGCTCTTTGGGAAGGGTGACT 362
    |||  |||  |||  |||  |||  ||  ||||||||  ||||||||||||||||||
194 CCACATCTTCACTACCAGGAGTCGTACCCGGCTTTTTGGGAAGGGTGACT 243

363 CGGAGGAGGCTTTCCCGGTGGATTGCCCTCACGAGGAAGGTGAGCTGGAC 412
    ||||||||||  |  ||  ||||  ||||||  |  |||||||||  |  ||||||
244 CGGAGGAGGCCTCTCCCCTGGACTGCCCTTATGAGGAAGGCGGGCTGGCT 293

413 TCCTGCCCGACCATCACAGTCAGCCCTGTTATCACCATCCAGAGGCCAGG 462
    ||||||||  |  |||||||  |||||||  ||||||  |  ||  ||||||||||||  ||
294 TCCTGCCCTATCATCACTGTCAGCTCTGTTCTAACTATCCAGAGGCCTGG 343

463 AGACGGCCCCACCGGTGCCAGGCTGCTGTCCCAGGACTCTGTCGCCGCCA 512
    ||  ||  ||  ||  |||  ||||||  |  ||||||||||||  |||  ||||||
344 GGATGGACCTGCCAGTGTCAGGCCGTCATCCCAGGACTCCGTCTCCGCTG 393

513 GCACCGAGAAGACCCTCAGGCTCTATGATCGCAGGAGTATCTTTGAAGCC 562
    |  |||||||  |||  |||||||||||||||||||||||  ||||||  ||  ||
394 G...TGAGAAGCCCCCGAGGCTCTATGATCGCAGGAGCATCTTCGATGCT 440

563 GTTGCTCAGAATAACTGCCAGGATCTGGAGAGCCTGCTGCTCTTCCTGCA 612
    ||  ||||||||  ||||||||||||||  ||||||||||||||||||  |||||||||
441 GTGGCTCAGAGTAACTGCCAGGAGCTGGAGAGCCTGCTGCCCTTCCTGCA 490

613 GAAGAGCAAGAAGCACCTCACAGACAACGAGTTCAAAGACCCTGAGACAG 662
    ||  ||||||||||||  |||  ||  ||||  ||||||||||||||||  ||||||||
491 GAGGAGCAAGAAGCGCCTGACTGACAGCGAGTTCAAAGACCCAGAGACAG 540

663 GGAAGACCTGTCTGCTGAAAGCCATGCTCAACCTGCACGACGGACAGAAC 712
    |  ||||||||||||||  ||||||||||||||  ||||||  |  ||  ||||||
541 GAAAGACCTGTCTGCTAAAAGCCATGCTCAATCTGCACAATGGGCAGAAT 590

713 ACCACCATCCCCCTGCTCCTGGAGATCGCGCGGCAAACGGACAGCCTGAA 762
    |||||||  |  ||||||||||||  |  ||  |||  |  ||  ||||||||||||
591 GACACCATCGCTCTGCTCCTGGACGTTGCCCGGAAGACAGACAGCCTGAA 640
```

FIG.2A

```
763 GGAGCTTGTCAACGCCAGCTACACGGACAGCTACTACAAGGGCCAGACAG 812
    | || ||||||| |||||||||| |||||||||||||||||||||||||||
641 GCAGTTTGTCAATGCCAGCTACACAGACAGCTACTACAAGGGCCAGACAG 690

813 CACTGCACATCGCCATCGAGAGACGCAACATGGCCCTGGTGACCCTCCTG 862
    |||||||||| ||||| ||  | || ||||||| | |||||||||||  ||
691 CACTGCACATTGCCATTGAACGGCGGAACATGACGCTGGTGACCCTCTTG 740

863 GTGGAGAACGGAGCAGACGTCCAGGCTGCGGCCCATGGGGACTTCTTTAA 912
    |||||||| |||||||| |||||||||||||| | |||||||||||  ||
741 GTGGAGAATGGAGCAGATGTCCAGGCTGCGGCTAACGGGGACTTCTTCAA 790

913 GAAAACCAAAGGGCGGCCTGGATTCTACTTCGGTGAACTGCCCCTGTCCC 962
    |||||||||||||| ||||||| |||||||| ||||| ||||||||||||
791 GAAAACCAAAGGGAGGCCTGGCTTCTACTTTGGTGAGCTGCCCCTGTCCC 840

963 TGGCCGCGTGCACCAACCAGCTGGGCATCGTGAAGTTCCTGCTGCAGAAC 1012
    |||| |||||||||||||||||| ||| ||||||||||||||||||||||
841 TGGCTGCGTGCACCAACCAGCTGGCCATTGTGAAGTTCCTGCTGCAGAAC 890

1013 TCCTGGCAGACGGCCGACATCAGCGCCAGGGACTCGGTGGGCAACACGGT 1062
     ||||||||| | || |||||||||||| ||||||| ||||||||||||||
891  TCCTGGCAGCCTGCAGACATCAGCGCCCGGGACTCAGTGGGCAACACGGT 940

1063 GCTGCACGCCCTGGTGGAGGTGGCCGACAACACGGCCGACAACACGAAGT 1112
     ||| || |||||||||||||||||| || ||||| | ||||||||| ||||
941  GCTTCATGCCCTGGTGGAGGTGGCAGATAACACAGTTGACAACACCAAGT 990

1113 TTGTGACGAGCATGTACAATGAGATTCTGATCCTGGGGGCCAAACTGCAC 1162
     | || || |||||||||||| ||||| ||||||||||||||||||| |||
991  TCGTGACAAGCATGTACAACGAGATCTTGATCCTGGGGGCCAAACTCCAC 1040

1163 CCGACGCTGAAGCTGGAGGAGCTCACCAACAAGAAGGGAATGACGCCGCT 1212
     || || |||||||||| ||| ||||||||| |||||| | ||||| ||
1041 CCCACGCTGAAGCTGGAAGAGATCACCAACAGGAAGGGGCTCACGCCACT 1090

1213 GGCTCTGGCAGCTGGGACCGGGAAGATCGGGGTCTTGGCCTATATTCTCC 1262
     ||||||||| ||| | |  |||||||||||||||||||||| |||||||
1091 GGCTCTGGCTGCTAGCAGTGGGAAGATCGGGGTCTTGGCCTACATTCTCC 1140

1263 AGCGGGAGATCCAGGAGCCCGAGTGCAGGCACCTGTCCAGGAAGTTCACC 1312
     || |||||||||| || ||||||||| | ||||| ||||||||||||||||
1141 AGAGGGAGATCCATGAACCCGAGTGCCGACACCTATCCAGGAAGTTCACC 1190
```

FIG.2B

```
1313 GAGTGGGCCTACGGGCCCGTGCACTCCTCGCTGTACGACCTGTCCTGCAT 1362
     || |||||||| ||||| |||||||||||| || || ||||||||||||||
1191 GAATGGGCCTATGGGCCAGTGCACTCCTCCCTTTATGACCTGTCCTGCAT 1240

1363 CGACACCTGCGAGAAGAACTCGGTGCTGGAGGTGATCGCCTACAGCAGCA 1412
     |||||||| || |||||||||||| ||||||||||||||| ||||||||||
1241 TGACACCTGTGAAAAGAACTCGGTTCTGGAGGTGATCGCTTACAGCAGCA 1290

1413 GCGAGACCCCTAATCGCCACGACATGCTCTTGGTGGAGCCGCTGAACCGA 1462
     | |||||||||||| || || |||||||| | ||||| || ||||||||
1291 GTGAGACCCCTAACCGTCATGACATGCTTCTCGTGGAACCCTTGAACCGA 1340

1463 CTCCTGCAGGACAAGTGGGACAGATTCGTCAAGCGCATCTTCTACTTCAA 1512
     ||||| ||||||||||||||||||||| ||||||||||||||||||||||
1341 CTCCTACAGGACAAGTGGGACAGATTTGTCAAGCGCATCTTCTACTTCAA 1390

1513 CTTCCTGGTCTACTGCCTGTACATGATCATCTTCACCATGGCTGCCTACT 1562
     |||| | ||||||||| |||| |||||||||||||| ||||||||||
1391 CTTCTTCGTCTACTGCTTGTATATGATCATCTTCACCGCGGCTGCCTACT 1440

1563 ACAGGCCCGTGGATGGCTTGCCTCCCTTTAAGATGGAAAAAAC...TGGA 1609
     |  |||| ||||| |||||||| |||| |||| || |||| ||   |||
1441 ATCGGCCTGTGGAAGGCTTGCCCCCCTATAAGCTGAAAAACACCGTTGGG 1490

1610 GACTATTTCCGAGTTACTGGAGAGATCCTGTCTGTGTTAGGAGGAGTCTA 1659
     |||||||||||||| || ||||||||| ||||||||| ||||||||||||
1491 GACTATTTCCGAGTCACCGGAGAGATCTTGTCTGTGTCAGGAGGAGTCTA 1540

1660 CTTCTTTTTCCGAGGGATTCAGTATTTCCTGCAGAGGCGGCCGTCGATGA 1709
     |||||| |||||||||||||| |||||||||||||||| || ||  | |
1541 CTTCTTCTTCCGAGGGATTCAATATTTCCTGCAGAGGCGACCATCCCTCA 1590

1710 AGACCCTGTTTGTGGACAGCTACAGTGAGATGCTTTTCTTTCTGCAGTCA 1759
     |||   |||||||||||||||||||||||| ||||||||| | |||||
1591 AGAGTTTGTTTGTGGACAGCTACAGTGAGATACTTTTCTTTGTACAGTCG 1640

1760 CTGTTCATGCTGGCCACCGTGGTGCTGTACTTCAGCCACCTCAAGGAGTA 1809
     |||||||||||||| | ||||| |||||||||||||| | |||||||||
1641 CTGTTCATGCTGGTGTCTGTGGTACTGTACTTCAGCCAACGCAAGGAGTA 1690

1810 TGTGGCTTCCATGGTATTCTCCCTGGCCTTGGGCTGGACCAACATGCTCT 1859
     |||||||||||||| |||||||||| |||||||||||||||||||||||
1691 TGTGGCTTCCATGGTGTTCTCCCTGGCCATGGGCTGGACCAACATGCTCT 1740
```

FIG.2C

```
1860 ACTACACCCGCGGTTTCCAGCAGATGGGCATCTATGCCGTCATGATTGAG 1909
     |||| ||||| || |||||||||||||||||||||||| ||||||||||||
1741 ACTATACCCGAGGATTCCAGCAGATGGGCATCTATGCTGTCATGATTGAG 1790

1910 AAGATGATCCTGAGAGACCTGTGCCGTTTCATGTTTGTCTACATCGTCTT 1959
     |||||||||||| |||||||||||||| || ||||| |||||| |||| ||
1791 AAGATGATCCTCAGAGACCTGTGCCGGTTTATGTTCGTCTACCTCGTGTT 1840

1960 CTTGTTCGGGTTTTCCACAGCGGTGGTGACGCTGATTGAAGACGGGAAGA 2009
     |||||| || |||||||||||| |||||||| |||||||| || ||||||||
1841 CTTGTTTGGATTTTCCACAGCTGTGGTGACACTGATTGAGGATGGGAAGA 1890

2010 ATGACTCCCTGCCGTCTGAGTCCACGTCGCACAGGTGGCGGGGGCCTGCC 2059
     || |||| |||||    |||||||| |||||| ||| |||||| |||||
1891 ATAACTCTCTGCCTATGGAGTCCACACCACACAAGTGCCGGGGGTCTGCC 1940

2060 TGCAGGCCCCCCGATAGCTCCTACAACAGCCTGTACTCCACCTGCCTGGA 2109
     |||| |   || | || ||| |||||||||||||| ||||| || |||||
1941 TGCAAG...CCAGGTAACTCTTACAACAGCCTGTATTCCACATGTCTGGA 1987

2110 GCTGTTCAAGTTCACCATCGGCATGGGCGACCTGGAGTTCACTGAGAACT 2159
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1988 GCTGTTCAAGTTCACCATCGGCATGGGCGACCTGGAGTTCACTGAGAACT 2037

2160 ATGACTTCAAGGCTGTCTTCATCATCCTGCTGCTGGCCTATGTAATTCTC 2209
     | |||||||||||||||||||||||||| | |||||||||| ||||||
2038 ACGACTTCAAGGCTGTCTTCATCATCCTGTTACTGGCCTATGTGATTCTC 2087

2210 ACCTACATCCTCCTGCTCAACATGCTCATCGCCCTCATGGGTGAGACTGT 2259
     |||||||||| |||||||||||||||| || |||||||||||||| ||
2088 ACCTACATCCTTCTGCTCAACATGCTCATTGCTCTCATGGGTGAGACCGT 2137

2260 CAACAAGATCGCACAGGAGAGCAAGAACATCTGGAAGCTGCAGAGAGCCA 2309
     ||||||||| ||||| ||||||||||||||||||||||||||||||||||
2138 CAACAAGATTGCACAAGAGAGCAAGAACATCTGGAAGCTGCAGAGAGCCA 2187

2310 TCACCATCCTGGACACGGAGAAGAGCTTCCTTAAGTGCATGAGGAAGGCC 2359
     |||||||||||||| || |||||||||||||| ||||||||||||||||||
2188 TCACCATCCTGGATACAGAGAAGAGCTTCCTGAAGTGCATGAGGAAGGCC 2237

2360 TTCCGCTCAGGCAAGCTGCTGCAGGTGGGGTACACACCTGATGGCAAGGA 2409
     |||||||| |||||||||||||||||||||| ||| ||||| |||||||||
2238 TTCCGCTCTGGCAAGCTGCTGCAGGTGGGGTTCACTCCTGACGGCAAGGA 2287
```

FIG.2D

```
2410 CGACTACCGGTGGTGCTTCAGGGTGGACGAGGTGAACTGGACCACCTGGA 2459
     ||||||||||||| |||||||||||||||||| |||||||| ||||||||
2288 TGACTACCGGTGGTGTTTCAGGGTGGACGAGGTAAACTGGACTACCTGGA 2337

2460 ACACCAACGTGGGCATCATCAACGAAGACCCGGGCAACTGTGAGGGCGTC 2509
     ||||||| ||||| |||||||||||| ||||| ||||||||||||||||||
2338 ACACCAATGTGGGTATCATCAACGAGGACCCAGGCAACTGTGAGGGCGTC 2387

2510 AAGCGCACCCTGAGCTTCTCCCTGCGGTCAAGCAGAGTTTCAGGCAGACA 2559
     |||||||||||||||||||||||||| ||||| || |||||||||| ||| |
2388 AAGCGCACCCTGAGCTTCTCCCTGAGGTCAGGCCGAGTTTCAGGGAGAAA 2437

2560 CTGGAAGAACTTTGCCCTGGTCCCCCTTTTAAGAGAGGCAAGTGCTCGAG 2609
     |||||||||||||||||||||| |||||| | || || ||||| ||||||
2438 CTGGAAGAACTTTGCCCTGGTTCCCCTTCTGAGGGATGCAAGCACTCGAG 2487

2610 ATAGGCAGTCTGCTCAGCCCGAGGAAGTTTATCTGCGACAGTTTTCAGGG 2659
     |||| || | | |||| || ||||||| | ||| || | | | ||
2488 ATAGACATGCCACCCAGCAGGAAGAAGTTCAACTGAAGCATTATACGGGA 2537

2660 TCTCTGAAGCCAGAGGACGCTGAGGTCTTCAAGAGTCCTGCCGCTTCCGG 2709
     || || |||||||||||| |||||||| |||||| | | | ||| ||
2538 TCCCTTAAGCCAGAGGATGCTGAGGTTTTCAAGGATTCCATGGTCCCAGG 2587

2710 GGAGAAGTGA.GGACGTCACGCAGACAGCACTGTCAACACTGGGCCTTAG 2758
     ||||||| | || |||| | |||| | |||| | ||| |||
2588 GGAGAAATAATGGACACTATGCAGGGA......TCAATGCGGGGTCTT.. 2629

2759 GAGACCCCGTTGCCACGGGGGGCTTGCTGAGGGAA.CACCAGTGCTCTGT 2807
                    ||| || |||| |||||| || |||| | ||
2630 ..............TGGGTGGTCTGCTTAGGGAACCAGCAGGG..TTGA 2662

2808 CAGCAGCCTGGCCTGGTCTGTGCCTGCCCA.GCATGTTCCCAAATCTGTG 2856
     |  | | || |  ||||||||||||| | ||| |||| |  || |
2663 CGTTATCTGGGTCCACTCTGTGCCTGCCTAGGCACATTCCTAGGACTTCG 2712

2857 CTGGACAAGCTGTGGGAA 2874
     || |  ||||||||||
2713 GCGGGCCTGCTGTGGGAA 2730
```

FIG.2E

```
           1                                                          60
hvr1  ---------- ---------- ---------- ------MKKW SSTDLGAAAD PLQKDTCPDP
rvr1  ---------- ---------- ---------- ------MEQR ASLDSEESES PPQENSCLDP
hvr2  ---------- ---------- ---------- ---------- ---------- ----------
hvr3  MADSSEGPRA GPGEVAELPG DESGTPGGEA FPLSSLANLF EGEDGSLSPS PADASRPAGP
Consensus ---------- ---------- ---------- ---------- ---------- ----------

           61                                                         120
hvr1  LDGDPNSRPP PAKPQLSTAK SRTRLFGKGD SEEAFPVDCP HEEGELD.SC PTITVSPVIT
rvr1  PDRDPNCKPP PVKPHIFTTR SRTRLFGKGD SEEASPLDCP YEEGGLA.SC PIITVSSVLT
hvr2  ---------- ------MTSP SSSPVFRLET LDGGQEDGSE ADRGKLDFGS GLPPMESQFQ
hvr3  GDGRPNLRMK FQGAFRKGVP NPIDLLESTL YESSVVPGPK KAPMDSLFDY GTYRHHSS.D
Consensus ---------- ---------- ---------- ---------- ---------- ----------

           121                                                        180
hvr1  IQRPGDGPTG ARLLSQDSVA ASTEKTLRLY DRRSIFEAVA QNNCQDLESL LLFLQKSKKH
rvr1  IQRPGDGPAS VRPSSQDSVS AG.EKPPRLY DRRSIFDAVA QSNCQELESL LPFLQRSKKR
hvr2  GEDRKFAPQI RVNLNYRKGT GASQPDPNRF DRDRLFNAVS RGVPEDLAGL PEYLSKTSKY
hvr3  NKRWRKKIIE KQPQSPKAPA PQPPPILKVF NRPILFDIVS RGSTADLDGL LPFLLTHKKR
Consensus ---------- ---------- ---------- -R---F--V- ------L--L ---L----K-

           181                                                        240
hvr1  LTDNEFKDPE TGKTCLLKAM LNLHDGQNTT IPLLLEIARQ TDSLKELVNA SYTDSYYKGQ
rvr1  LTDSEFKDPE TGKTCLLKAM LNLHNGQNDT IALLLDVARK TDSLKQFVNA SYTDSYYKGQ
hvr2  LTDSEYTEGS TGKTCLMKAV LNLKDGVNAC ILPLLQIDRD SGNPQPLVNA QCTDDYYRGH
hvr3  LTDEEFREPS TGKTCLPKAL LNLSNGRNDT IPVLLDIAER TGNMREFINS PFRDIYYRGQ
Consensus LTD-E----- TGKTCL-KA- LNL--G-N-- I--LL----- --------N- ---D-YY-G-

           241                                                        300
hvr1  TALHIAIERR NMALVTLLVE NGADVQAAAH GDFFKKTKGR PGFYFGELPL SLAACTNQLG
rvr1  TALHIAIERR NMTLVTLLVE NGADVQAAAN GDFFKKTKGR PGFYFGELPL SLAACTNQLA
hvr2  SALHIAIEKR SLQCVKLLVE NGANVHARAC GRFFQKGQG. TCFYFGELPL SLAACTKQWD
hvr3  TALHIAIERR CKHYVELLVA QGADVHAQAR GRFFQPKDEG GYFYFGELPL SLAACTNQPH
Consensus -ALHIAIE-R ----V-LLV- -GA-V-A-A- G-FF------ --FYFGELPL SLAACT-Q--

           301                                                        360
hvr1  IVKFLLQNSW QTADISARDS VGNTVLHALV EVADNTADNT KFVTSMYNEI LILGAKLHPT
rvr1  IVKFLLQNSW QPADISARDS VGNTVLHALV EVADNTVDNT KFVTSMYNEI LILGAKLHPT
hvr2  VVSYLLENPH QPASLQATDS QGNTVLHALV MISDNSAENI ALVTSMYDGL LQAGARLCPT
hvr3  IVNYLTENPH KKADMRRQDS RGNTVLHALV AIADNTRENT KFVTKMYDLL LLKCARLFPD
Consensus -V--L--N-- --A-----DS -GNTVLHALV ---DN---N- --VT-MY--- L---A-L-P-

           361                                                        420
hvr1  LKLEELTNKK GMTPLALAAG TGKIGVLAYI LQREIQEPEC RHLSRKFTEW AYGPVHSSLY
rvr1  LKLEEITNRK GLTPLALAAS SGKIGVLAYI LQREIHEPEC RHLSRKFTEW AYGPVHSSLY
hvr2  VQLEDIRNLQ DLTPLKLAAK EGKIEIFRHI LQREFS..GL SHLSRKFTEW CYGPVRVSLY
hvr3  SNLEAVLNND GLSPLMMAAK TGKIGIFQHI IRREVTDEDT RHLSRKFKDW AYGPVYSSLY
Consensus --LE---N-- ---PL--AA- -GKI-----I -RE------- -HLSRKF--W -YGPV--SLY

           421                                                        480
hvr1  DLSCIDTC.E KNSVLEVIAY SSSETPNRHD MLLVEPLNRL LQDKWDRFVK RIFYFNFLVY
rvr1  DLSCIDTC.E KNSVLEVIAY SSSETPNRHD MLLVEPLNRL LQDKWDRFVK RIFYFNFFVY
hvr2  DLASVDSC.E ENSVLEIIAF .HCKSPHRHR MVVLEPLNKL LQAKWDLLIP K.FFLNFLCN
hvr3  DLSSLDTCGE EASVLEILVY .NSKIENRHE MLAVEPINEL LRDKWRKFGA VSFYINVVSY
Consensus DL---D-C-E --SVLE---- -------RH- M---EP-N-L L--KW----- --F--N----
```

FIG.3A

```
           481                                                       540
hvr1  CLYMIIFTMA AYYRP.VDGL PPFKMEKT.G DYFRVTGEIL SVLGGVYFFF RGIQ.YFLQR
rvr1  CLYMIIFTAA AYYRP.VEGL PPYKLKNTVG DYFRVTGEIL SVSGGVYFFF RGIQ.YFLQR
hvr2  LIYMFIFTAV AYHQPTLKKQ AAPHLKAEVG NSMLLTGHIL ILLGGIYLLV GQLW.YFWRR
hvr3  LCAMVIFTLT AYYQP.LEGT PPYPYRTTV. DYLRLAGEVI TLFTGVLFFF TNIKDLFMKK
Consensus ---M-IFT-- AY--P----- ---------- ------G--- ----G----- ------F---

           541                                                       600
hvr1  RPSMKTLFVD SYSEMLFFLQ SLFMLATVVL YFSHLKEYVA SMVFSLALGW TNMLYYTRGF
rvr1  RPSLKSLFVD SYSEILFFVQ SLFMLVSVVL YFSQRKEYVA SMVFSLAMGW TNMLYYTRGF
hvr2  HVFIWISFID SYFEILFLFQ ALLTVVSQVL CFLAIEWYLP LLVSALVLGW LNLLYYTRGF
hvr3  CPGVNSLFID GSFQLLYFIY SVLVIVSAAL YLAGIEAYLA VMVFALVLGW MNALYFTRGL
Consensus -------F-D -----L---- ---------L -------Y-- --V--L--GW -N-LY-TRG-

           601                                                       660
hvr1  QQMGIYAVMI EKMILRDLCR FMFVYIVFLF GFSTAVVTLI EDGKNDSLP. ...SESTSHR
rvr1  QQMGIYAVMI EKMILRDLCR FMFVYLVFLF GFSTAVVTLI EDGKNNSLP. ...MESTPHK
hvr2  QHTGIYSVMI QKVILRDLLR FLLIYLVFLF GFAVALVSLS QEAWRPEAPT GPNATESVQP
hvr3  KLTGTYSIMI QKILFKDLFR FLLVYLLFMI GYASALVSLL NPCANMKVCN EDQTNCTVPT
Consensus ---G-Y--MI -K----DL-R F---Y--F-- G---A-V-L- ---------- ----------

           661                                                       720
hvr1  WRGPACRPPD SSYNSLYSTC LELFKFTIGM GQLEFTENYD FKAVFIILLL AYVILTYILL
rvr1  CRGSACK.PG NSYNSLYSTC LELFKFTIGM GQLEFTENYD FKAVFIILLL AYVILTYILL
hvr2  MEGQEDEGNG AQYRGILEAS LELFKFTIGM GELAFQEQLH FRGMVLLLLL AYVLLTYILL
hvr3  YPSCRDSETF STF......L LDLFKLTIGM GQLEMLSSTK YPVVFIILLV TYIILTFVLL
Consensus ---------- ---------- L-LFK-TIGM G-L------- -------LL- -Y--LT--LL

           721                                                       780
hvr1  LNMLIALMGE TVNKIAQESK NIWKLQRAIT ILDTEKSFLK CMRKAFRSGK LLQVGYTPDG
rvr1  LNMLIALMGE TVNKIAQESK NIWKLQRAIT ILDTEKSFLK CMRKAFRSGK LLQVGFTPDG
hvr2  LNMLIALMSE TVNSVATDSW SIWKLQKAIS VLXMENGYWW C.RKKQRAGV MLTVGTKPDG
hvr3  LNMLIALMGE TVGQVSKESK HIWKLQWATT ILDIERSFPV FLRKAFRSGE MVTVGKSSDG
Consensus LNMLIALM-E TV------S- -IWKLQ-A-- -L--E----- --RK--R-G- ---VG---DG

           781                                                       840
hvr1  KDDYRWCFRV DEVNWTTWNT NVGIINEDPG NCEGVKRTLS FSLRSS.... RVSGRHWKNF
rvr1  KDDYRWCFRV DEVNWTTWNT NVGIINEDPG NCEGVKRTLS FSLRSG.... RVSGRNWKNF
hvr2  SPDERWCFRV EEVNWASWEQ TLPTLCEDPS GA.GVPRTLE NPVLASPPKE DEDGASEENY
hvr3  TPDRRWCFRV DEVNWSHWNQ NLGIINEDPG KNE....TYQ YYGFSHTVGR LRRDRWSSVV
Consensus --D-RWCFRV -EVNW--W-- ------EDP- -------T-- ---------- ----------

           841                                                  889
hvr1  ALVPLLREAS ARDRQSAQPE EVYLRQFSGS LKPEDAEVFK SPAASGEK*
rvr1  ALVPLLPDAS TRDRHATQQE EVQLKHYTGS LKPEDAEVFK DSMVPGEK~
hvr2  VPVQLLQSN* ---------- ---------- ---------- ----------
hvr3  PRVVELNKNS NPDEVVVPLD SMGNPRCDGH QQGYPRKWRT DDAPL*~~~
Consensus --V--L---- ---------- ---------- ---------- ---------
```

FIG.3B

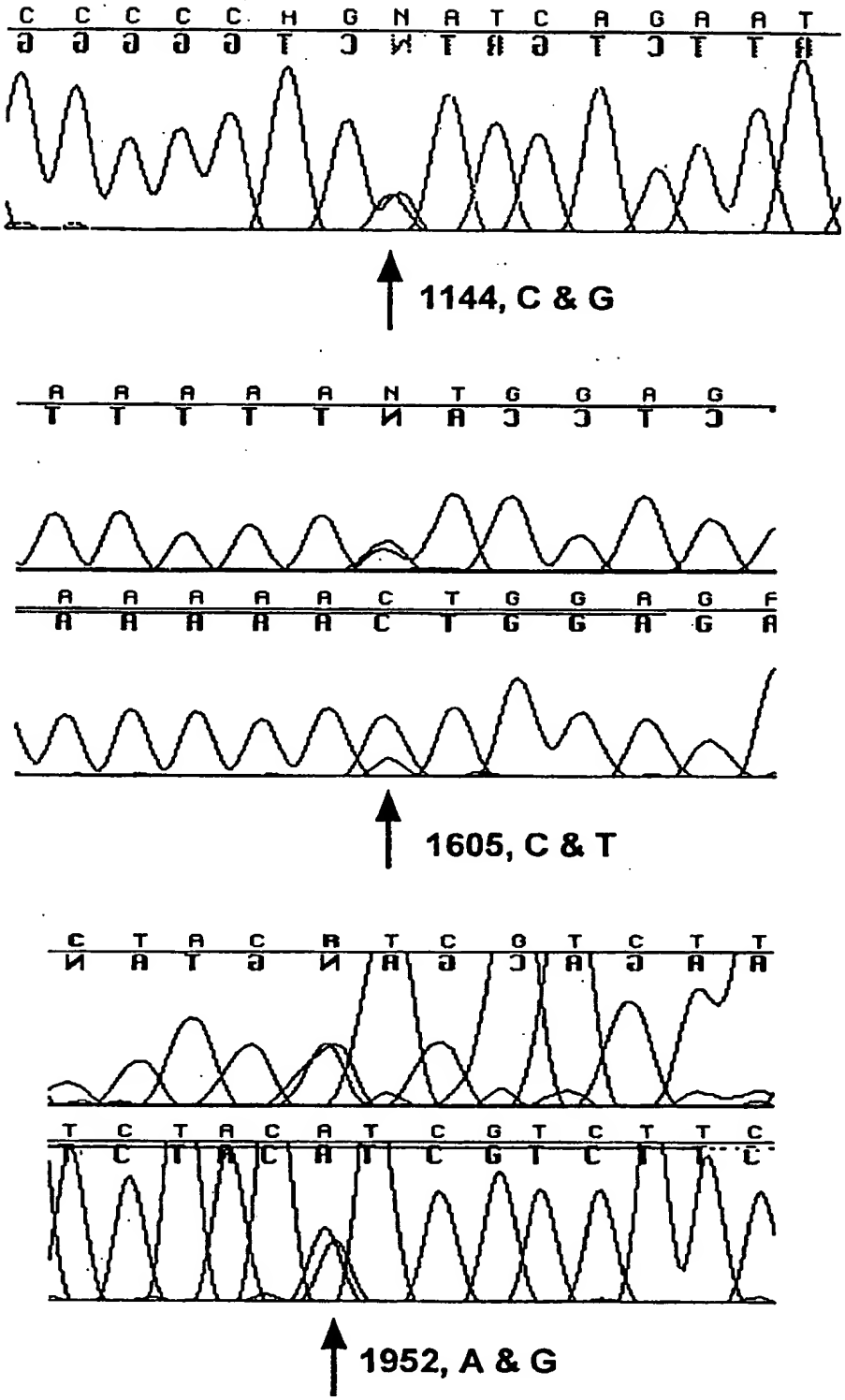

FIG.4

```
301 CTGAAATGTTAGGCAGATACCCTGTGGGAGCTTTGTTCTGGGATGCTAAGAACCGCTTGA 360
  1 .........GAGGTTCAGTCCTGGAAACACTTCAGTTCTAGGGGGCTGGGGGCAGCAGCA  51
361 GGATTTAAGCTTTGCCACTTTGGCTCCGGAGCAAGGGCAGAGGGCTGAGCAGTGCAGACG 420
 52 AGTTGGAGTTTTGGGGTACCCTGCTTC....ACAGGGCCCTTGGCAAGGAGGGCAGGTGG 107
421 GGCCTGGGGCAGGCATGGCGGATTCCAGCGAAGGCCCCCGCGCGGGGCCCGGGGAGGTGG 480
108 GGTCTAAGGACAAGCAGTCCTTACTTTGGGAGTCAACCCCGGCGTGGTGGCTGCTGCAGG 167
481 CTGAGCTCCCCGGGGATGAGAGTGGCACCCCAGGTGGGGAGGCTTTTCCTCTCTCCTCCC 540
168 TT..GCACACTGGGCCACAGAGGATCCAGCAAGGATGAAGAAATGGAGCAGCACAGACTT 225
541 TGGCCAATCTGTTTGAGGGGGAGGATGGCTCCCTTTCGCCCTCACCGGCTGATGCCAGTC 600
226 GGGGGCAGCTGCGGACCCACTCCAAAAGGACACCTGCCCAGACCCCCTGGATGGAGACCC 285
601 GCCCTGCTGGCCCAGGCGATGGGCGACCAAATCTGCGCATGAAGTTCCAGGGCGCCTTCC 660
286 TAACTCCAGGCC.....ACCTCCAGCCAAGCCCCAGCTCTCCACGG 326
661 GCAAGGGGGTGCCCAACCCCATCGATCTGCTGGAGTCCACCCTATATGAGTCCTCGGTGG 720
327 CCAAGAG...CCGCACCCGGCTCTTTGGGAAGGGTGACTCGGAGGAGGCTTTCCCGGTGG 383
721 TGCCTGGGCCCAAGAAAGCACCCATGGACTCACTGTTTGACTACGGCACCTATCGTCACC 780
384 ATTGCCCTCACGAGGAAGGTGAGCTGGACTC.CTGCCCGACCATCACAGTCA.....GCC 437
781 ACTCCAGTGACAACAAGAGGTGGAGGAAGAAGATCATAGAGAAGCAGCCGCAGAGCCCCA 840
438 CTGTTATCACCATCCAGAGGCCAGGAGACGGCCCCACCGGTGCCAGGCTGCTGTCCCAGG 497
841 AAGCCCCTGCCCCTCAGCCGCCCCCCATCCTCAAAGTCTTCAACCGGCCTATCCTCTTTG 900
498 ACTCTGTCGCCGCCAGCACCGAGAAGACCCTCAGGCTCTATGATCGCAGGAGTATCTTTG 557
901 ACATCGTGTCCCGGGGCTCCACTGCTGACCTGGACGGGCTGCTCCCATTCTTGCTGACCC 960
558 AAGCCGTTGCTCAGAATAACTGCCAGGATCTGGAGAGCCTGCTGCTCTTCCTGCAGAAGA 617
961 ACAAGAAACGCCTAACTGATGAGGAGTTTCGAGAGCCATCTACGGGGAAGACCTGCCTGC 1020
618 GCAAGAAGCACCTCACAGACAACGAGTTCAAAGACCCTGAGACAGGGAAGACCTGTCTGC 677
1021 CCAAGGCCTTGCTGAACCTGAGCAATGGCCGCAACGACACCATCCCTGTGCTGCTGGACA 1080
678 TGAAAGCCATGCTCAACCTGCACGACGGACAGAACACCACCATCCCCCTGCTCCTGGAGA 737
```

FIG.5A

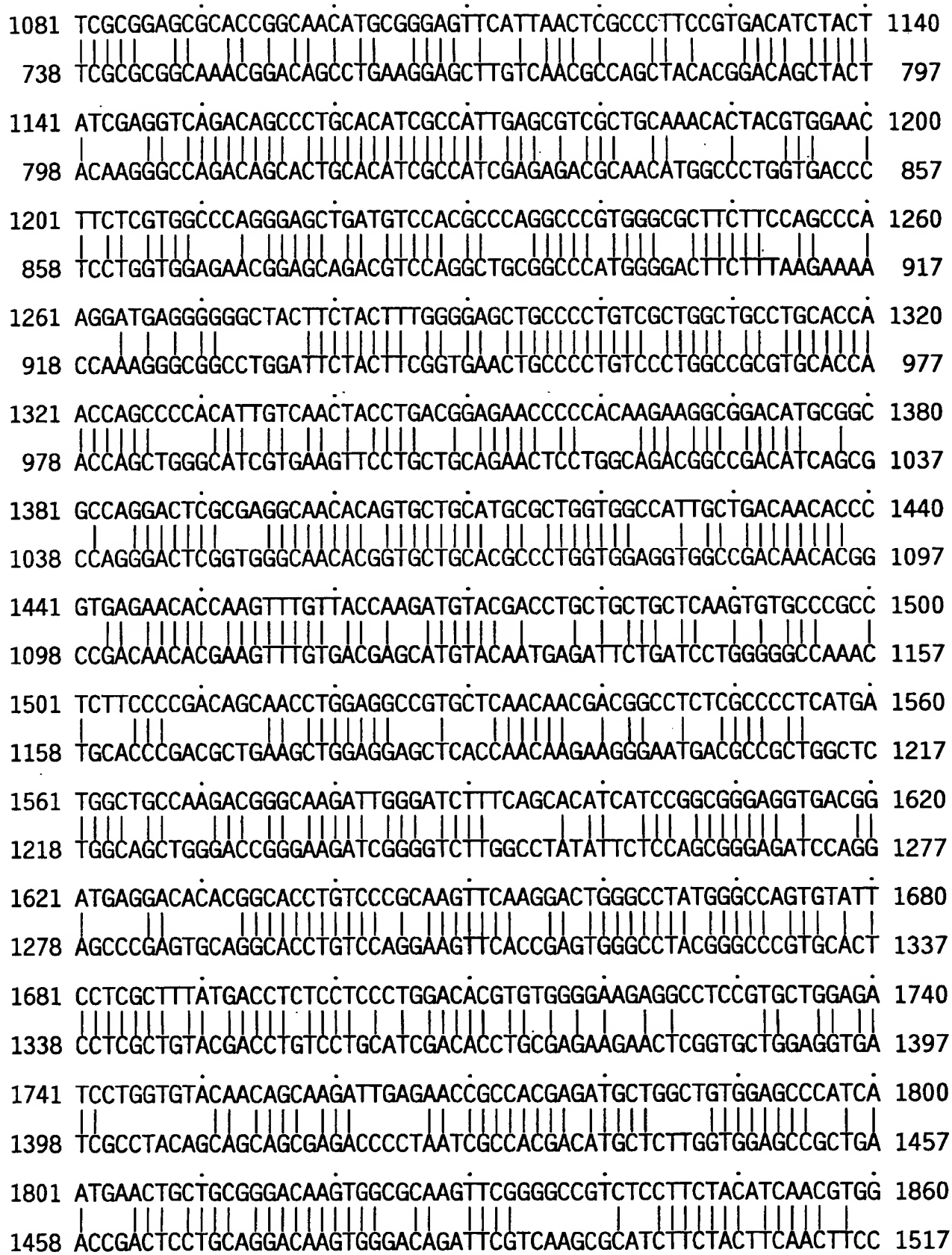
1081 TCGCGGAGCGCACCGGCAACATGCGGGAGTTCATTAACTCGCCCTTCCGTGACATCTACT 1140
738 TCGCGCGGCAAACGGACAGCCTGAAGGAGCTTGTCAACGCCAGCTACACGGACAGCTACT 797
1141 ATCGAGGTCAGACAGCCCTGCACATCGCCATTGAGCGTCGCTGCAAACACTACGTGGAAC 1200
798 ACAAGGGCCAGACAGCACTGCACATCGCCATCGAGAGACGCAACATGGCCCTGGTGACCC 857
1201 TTCTCGTGGCCCAGGGAGCTGATGTCCACGCCCAGGCCCGTGGGCGCTTCTTCCAGCCCA 1260
858 TCCTGGTGGAGAACGGAGCAGACGTCCAGGCTGCGGCCCATGGGGACTTCTTTAAGAAAA 917
1261 AGGATGAGGGGGGCTACTTCTACTTTGGGGAGCTGCCCCTGTCGCTGGCTGCCTGCACCA 1320
918 CCAAAGGGCGGCCTGGATTCTACTTCGGTGAACTGCCCCTGTCCCTGGCCGCGTGCACCA 977
1321 ACCAGCCCCACATTGTCAACTACCTGACGGAGAACCCCCACAAGAAGGCGGACATGCGGC 1380
978 ACCAGCTGGGCATCGTGAAGTTCCTGCTGCAGAACTCCTGGCAGACGGCCGACATCAGCG 1037
1381 GCCAGGACTCGCGAGGCAACACAGTGCTGCATGCGCTGGTGGCCATTGCTGACAACACCC 1440
1038 CCAGGGACTCGGTGGGCAACACGGTGCTGCACGCCCTGGTGGAGGTGGCCGACAACACGG 1097
1441 GTGAGAACACCAAGTTTGTTACCAAGATGTACGACCTGCTGCTGCTCAAGTGTGCCCGCC 1500
1098 CCGACAACACGAAGTTTGTGACGAGCATGTACAATGAGATTCTGATCCTGGGGGCCAAAC 1157
1501 TCTTCCCCGACAGCAACCTGGAGGCCGTGCTCAACAACGACGGCCTCTCGCCCCTCATGA 1560
1158 TGCACCCGACGCTGAAGCTGGAGGAGCTCACCAACAAGAAGGGAATGACGCCGCTGGCTC 1217
1561 TGGCTGCCAAGACGGGCAAGATTGGGATCTTTCAGCACATCATCCGGCGGGAGGTGACGG 1620
1218 TGGCAGCTGGGACCGGGAAGATCGGGGTCTTGGCCTATATTCTCCAGCGGGAGATCCAGG 1277
1621 ATGAGGACACACGGCACCTGTCCCGCAAGTTCAAGGACTGGGCCTATGGGCCAGTGTATT 1680
1278 AGCCCGAGTGCAGGCACCTGTCCAGGAAGTTCACCGAGTGGGCCTACGGGCCCGTGCACT 1337
1681 CCTCGCTTTATGACCTCTCCTCCCTGGACACGTGTGGGGAAGAGGCCTCCGTGCTGGAGA 1740
1338 CCTCGCTGTACGACCTGTCCTGCATCGACACCTGCGAGAAGAACTCGGTGCTGGAGGTGA 1397
1741 TCCTGGTGTACAACAGCAAGATTGAGAACCGCCACGAGATGCTGGCTGTGGAGCCCATCA 1800
1398 TCGCCTACAGCAGCAGCGAGACCCCTAATCGCCACGACATGCTCTTGGTGGAGCCGCTGA 1457
1801 ATGAACTGCTGCGGGACAAGTGGCGCAAGTTCGGGGCCGTCTCCTTCTACATCAACGTGG 1860
1458 ACCGACTCCTGCAGGACAAGTGGGACAGATTCGTCAAGCGCATCTTCTACTTCAACTTCC 1517

FIG.5B

```
1861 TCTCCTACCTGTGTGCCATGGTCATCTTCACTCTCACCGCCTACTACCAGCCGCTGGAGG 1920
     |   ||||         ||||  ||||||||||  |  | |||||||||  |||  |||| |
1518 TGGTCTACTGCCTGTACATGATCATCTTCACCATGGCTGCCTACTACAGGCCCGTGGATG 1577

1921 GCACACCGCCGTACCCTTACCGCACCACGGTGGACTACCTGCGGCTGGCTGGCGAGGTCA 1980
     ||   || || |           |  || |  |||||  | ||  |  |||| ||| ||
1578 GCTTGCCTCCCTTTAAGATGGAAAAAACTGGAGACTATTTCCGAGTTACTGGAGAGATCC 1637

1981 TTACGCTCTTCACTGGGGTCCTGTTCTTCTTCACCAACATCAAAGACTTGTTCATGAAGA 2040
     |     | ||    || |||   ||||| |||       ||      | |  ||| || |||
1638 TGTCTGTGTTAGGAGGAGTCTACTTCTTTTTCCGAGGGAT...TCAGTATTTCCTGCAGA 1694

2041 AATGCCCTGGAGTGAATTCTCTCTTCATTGATGGCTCCTTCCAGCTGCTCTACTTCATCT 2100
       | ||      ||||  | || ||  | ||  ||| |     || |||| | |||  |
1695 GGCGGCCGTCGATGAAGACCCTGTTTGTGGACAGCTACAGTGAGATGCTTTTCTTTCTGC 1754

2101 ACTCTGTCCTGGTGATCGTCTCAGCAGCCCTCTACCTGGCAGGGATCGAGGCCTACCTGG 2160
     | ||  |  |  || | | | | |  || ||| |         || |||  ||  |||
1755 AGTCACTGTTCATGCTGGCCACCGTGGTGCTGTACTTCAGCCACCTCAAGGAGTATGTGG 1814

2161 CCGTGATGGTCTTTGCCCTGGTCCTGGGCTGGATGAATGCCCTTTACTTCACCCGTGGGC 2220
     |   ||||| ||  |||||| | |||||||||  ||    || |||| |||||| ||
1815 CTTCCATGGTATTCTCCCTGGCCTTGGGCTGGACCAACATGCTCTACTACACCCGCGGTT 1874

2221 TGAAGCTGACGGGGACCTATAGCATCATGATCCAGAAGATTCTCTTCAAGGACCTTTTCC 2280
     | ||| || ||| | ||||  | |||||||  |||||||  || | |  ||||| | ||
1875 TCCAGCAGATGGGCATCTATGCCGTCATGATTGAGAAGATGATCCTGAGAGACCTGTGCC 1934

2281 GATTCCTGCTCGTCTACTTGCTCTTCATGATCGGCTACGCTTCAGCCCTGGTCTCCCTCC 2340
     | ||| || | |||||| | |||||  || |||| |   | ||||  |||| | ||
1935 GTTTCATGTTTGTCTACATCGTCTTCTTGTTCGGGTTTTCCACAGCGGTGGTGACGCTGA 1994

2341 TGAACCCGTGTGCCAACATGAAGGTGTGCAATGAGGACCAGACCAACTGCACAGTGCCCA 2400
     | |     |    ||         ||     ||||  | | |  || |       |
1995 TTGAAGACGGGAAGAATGACTCCCTGCCGTCTGAGTCCACGTCGCACAGGTGGCGGGGGC 2054

2401 CTTACCCCTCGTGCCGTGACAG......CGAGACCTTCAGCACCTTCCTCCTGGACCTGT 2454
     || |   | | ||  || ||       | | | | |   | ||  |  |||||| ||||
2055 CTGCCTGCAGGCCCCCCGATAGCTCCTACAACAGCCTGTACTCCACCTGCCTGGAGCTGT 2114
```

FIG.5C

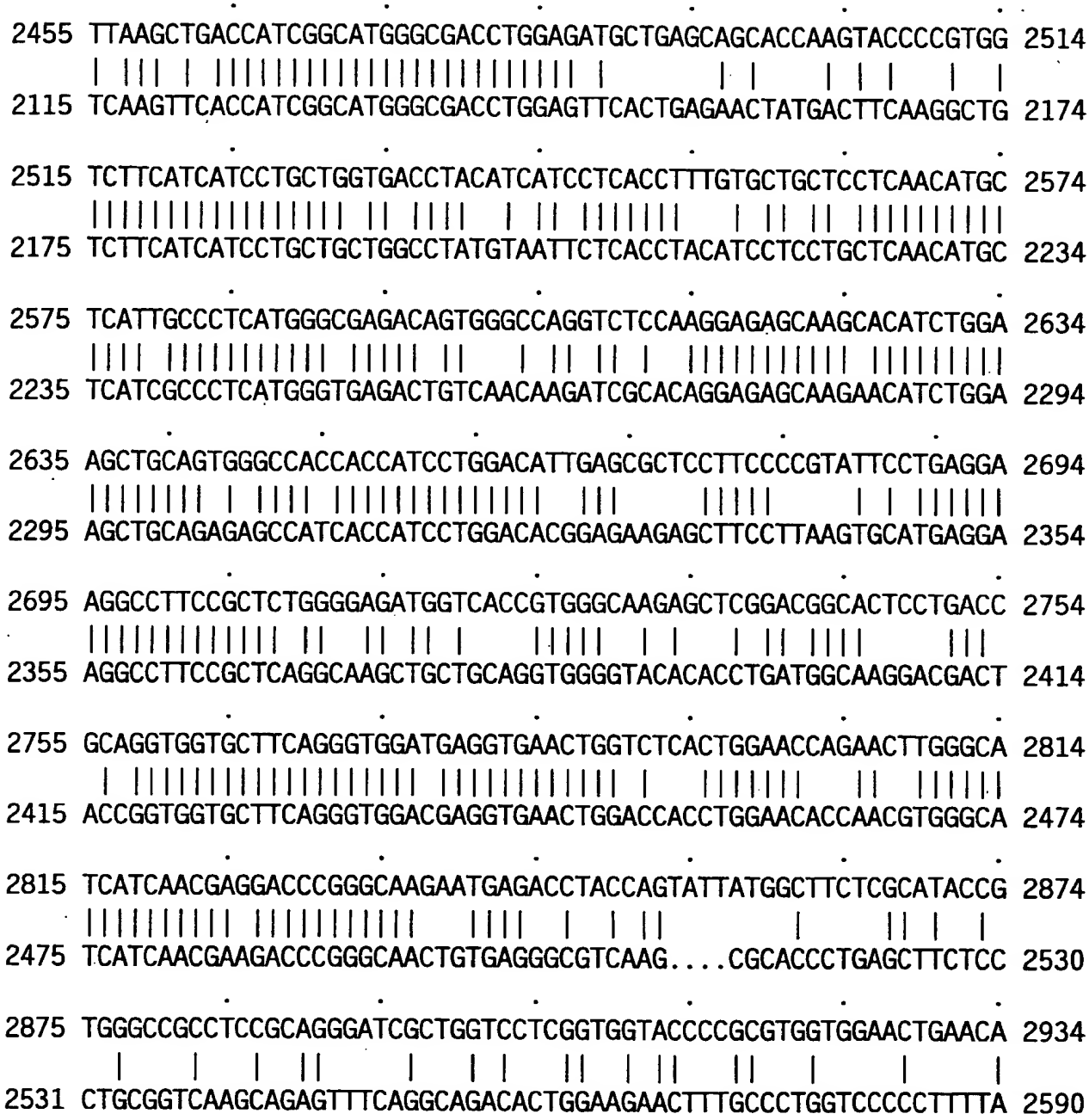
2455 TTAAGCTGACCATCGGCATGGGCGACCTGGAGATGCTGAGCAGCACCAAGTACCCCGTGG 2514
2115 TCAAGTTCACCATCGGCATGGGCGACCTGGAGTTCACTGAGAACTATGACTTCAAGGCTG 2174
2515 TCTTCATCATCCTGCTGGTGACCTACATCATCCTCACCTTTGTGCTGCTCCTCAACATGC 2574
2175 TCTTCATCATCCTGCTGCTGGCCTATGTAATTCTCACCTACATCCTCCTGCTCAACATGC 2234
2575 TCATTGCCCTCATGGGCGAGACAGTGGGCCAGGTCTCCAAGGAGAGCAAGCACATCTGGA 2634
2235 TCATCGCCCTCATGGGTGAGACTGTCAACAAGATCGCACAGGAGAGCAAGAACATCTGGA 2294
2635 AGCTGCAGTGGGCCACCACCATCCTGGACATTGAGCGCTCCTTCCCCGTATTCCTGAGGA 2694
2295 AGCTGCAGAGAGCCATCACCATCCTGGACACGGAGAAGAGCTTCCTTAAGTGCATGAGGA 2354
2695 AGGCCTTCCGCTCTGGGGAGATGGTCACCGTGGGCAAGAGCTCGGACGGCACTCCTGACC 2754
2355 AGGCCTTCCGCTCAGGCAAGCTGCTGCAGGTGGGGTACACACCTGATGGCAAGGACGACT 2414
2755 GCAGGTGGTGCTTCAGGGTGGATGAGGTGAACTGGTCTCACTGGAACCAGAACTTGGGCA 2814
2415 ACCGGTGGTGCTTCAGGGTGGACGAGGTGAACTGGACCACCTGGAACACCAACGTGGGCA 2474
2815 TCATCAACGAGGACCCGGGCAAGAATGAGACCTACCAGTATTATGGCTTCTCGCATACCG 2874
2475 TCATCAACGAAGACCCGGGCAACTGTGAGGGCGTCAAG....CGCACCCTGAGCTTCTCC 2530
2875 TGGGCCGCCTCCGCAGGGATCGCTGGTCCTCGGTGGTACCCCGCGTGGTGGAACTGAACA 2934
2531 CTGCGGTCAAGCAGAGTTTCAGGCAGACACTGGAAGAACTTTGCCCTGGTCCCCCTTTTA 2590

FIG.5D

```
2935 AGAACTCGAACCCGGACGAGGTGGTGGTGCCTCTGGACAGCATGGGGAACCCCCGCTGCG 2994
     |||    ||  |  |     | | | |    ||||  ||||  | ||||      |||||
2591 AGA....GAGGCAAGTGCTCGAGATAGGCAGTCTGCTCAGCCCGAGGAAGTTTATCTGCG 2646

2995 ATGGCCACCAGCAGGGTTACCCCCGCAAGTGGAGGACTGATGACGCCCCGCTCTAGGGAC 3054
     |   |      ||||||  |    || || ||| |          |     ||  |   |
2647 A...CAGTTTTCAGGGTCTCTGAAGCCAGAGGACGCTGAGGTCTTCAAGAGTCCTGCCGC 2703

3055 TGCAGCCCAGCCCCAGCTTCTCTGCCCACTCATTTCTAGTCCAGCCGCATTTCAGCAGTG 3114
     | | |   ||          |    | ||  ||   ||       | |    | || ||
2704 TTCCGGGGAGAAGTGAGGACGTCACGCAGACAGCACTGTCAACACTGGGCCTTAGGAGAC 2763

3115 CCTTCTGGGGTGTCCCCCCACACCCTGCTTTGGCCCCAGAGGCGAGGGACCAGTGGAGGT 3174
     ||        ||  || |      | ||||  ||   ||   | |      |||  |
2764 CC......CGTTGCCACGGGGGGCTTGCTGAGGGAACACCAGTGCTCTGTCAGCAGCCTG 2817

3175 GCCAGGGAGGCCCCAGGACCCTGTGGTCCCCTGGCTCTGCCTCCCCA.CCCTGGGGTGGG 3233
     ||| ||   |  || |  |    || ||||    || |||    | | |  ||||  | |
2818 GCCTGGTCTGTGCCTGCCCAGCATGTTCCCAAATCTGTGCTGGACAAGCTGTGGGAAGCG 2877

3234 GGCTCCCGGCCACCTGTCTTGCTCCTATGGA...GTCACATAAGCCAACGCCAGAGCCCC 3290
       ||      ||   |  || |  ||  |   |||||     | |  |      |  |
2878 TTCTTGGAAGCATGGGGAGTGATGACATCCAACCGTCACTGTCCCAAGTGAATCTCCTAC 2937

3291 TCCACCTCAGGCCCCAGCCCCTGCCTCTCCATTATTTATTTGCTCTGCTCTCAGGAAGCG 3350
          |||||    |   | | |    |  ||   |    ||||| ||   |      |
2938 AGACTTTCAGGTTTTACTACTTACTAAACAGTTGGATGGCAGCTCTACTGGGACATGTAG 2997

335I ACGTGACCCCTGCCCCAGCTGGAACCTGGCAGAGGCCTTAGGACCCCGTTCCAAGTGCAC 3410
     | |     ||     |         ||  ||     |  |||  |    |  ||
2998 GCCTTGGTTCTTGATTAATCTTTTTTTGAACGAATTACTTTTACCAGCTGATGATGGC.. 3055
```

FIG.5E

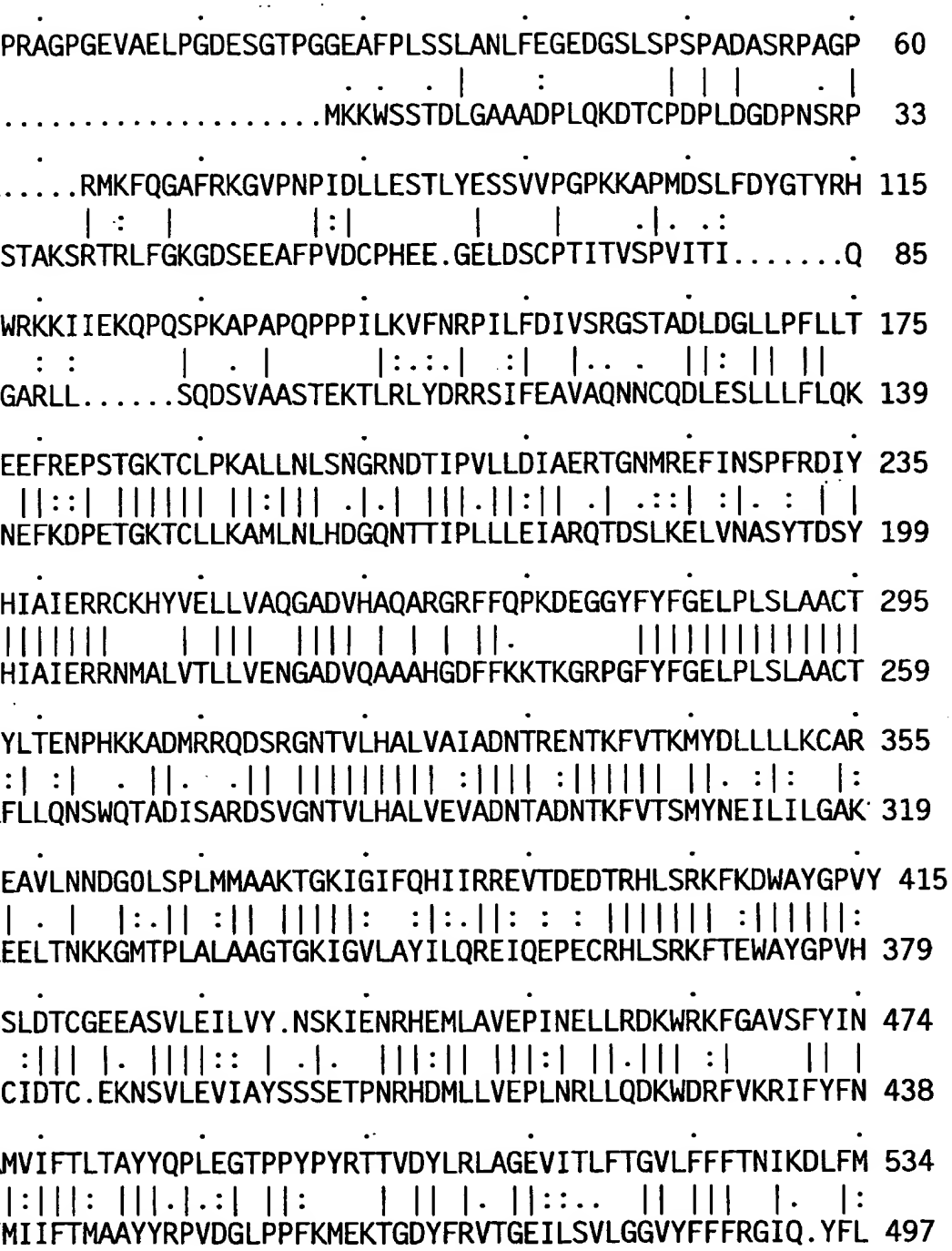
1 MADSSEGPRAGPGEVAELPGDESGTPGGEAFPLSSLANLFEGEDGSLSPSPADASRPAGP 60
1 .....MKKWSSTDLGAAADPLQKDTCPDPLDGDPNSRP 33
61 GDGRPNL.....RMKFQGAFRKGVPNPIDLLESTLYESSVVPGPKKAPMDSLFDYGTYRH 115
34 PPAKPQLSTAKSRTRLFGKGDSEEAFPVDCPHEE.GELDSCPTITVSPVITI.....Q 85
116 HSSDNKRWRKKIIEKQPQSPKAPAPQPPPILKVFNRPILFDIVSRGSTADLDGLLPFLLT 175
86 RPGDGPTGARLL.....SQDSVAASTEKTLRLYDRRSIFEAVAQNNCQDLESLLLFLQK 139
176 HKKRLTDEEFREPSTGKTCLPKALLNLSNGRNDTIPVLLDIAERTGNMREFINSPFRDIY 235
140 SKKHLTDNEFKDPETGKTCLLKAMLNLHDGQNTTIPLLLEIARQTDSLKELVNASYTDSY 199
236 YRGQTALHIAIERRCKHYVELLVAQGADVHAQARGRFFQPKDEGGYFYFGELPLSLAACT 295
200 YKGQTALHIAIERRNMALVTLLVENGADVQAAAHGDFFKKTKGRPGFYFGELPLSLAACT 259
296 NQPHIVNYLTENPHKKADMRRQDSRGNTVLHALVAIADNTRENTKFVTKMYDLLLLKCAR 355
260 NQLGIVKFLLQNSWQTADISARDSVGNTVLHALVEVADNTADNTKFVTSMYNEILILGAK 319
356 LFPDSNLEAVLNNDGQLSPLMMAAKTGKIGIFQHIIRREVTDEDTRHLSRKFKDWAYGPVY 415
320 LHPTLKLEELTNKKGMTPLALAAGTGKIGVLAYILQREIQEPECRHLSRKFTEWAYGPVH 379
416 SSLYDLSSLDTCGEEASVLEILVY.NSKIENRHEMLAVEPINELLRDKWRKFGAVSFYIN 474
380 SSLYDLSCIDTC.EKNSVLEVIAYSSSETPNRHDMLLVEPLNRLLQDKWDRFVKRIFYFN 438
475 VVSYLCAMVIFTLTAYYQPLEGTPPYPYRTTVDYLRLAGEVITLFTGVLFFFTNIKDLFM 534
439 FLVYCLYMIIFTMAAYYRPVDGLPPFKMEKTGDYFRVTGEILSVLGGVYFFFRGIQ.YFL 497

FIG.6A

```
535 KKCPGVNSLFIDGSFQLLYFIYSVLVIVSAALYLAGIEAYLAVMVFALVLGWMNALYFTR 594
498 QRRPSMKTLFVDSYSEMLFFLQSLFMLATVVLYFSHLKEYVASMVFSLALGWTNMLYYTR 557

595 GLKLTGTYSIMIQKILFKDLFRFLLVYLLFMIGYASALVSLLNPCANMKVCNEDQTNCTV 654
558 GFQQMGIYAVMIEKMILRDLCRFMFVYIVFLFGFSTAVVTLIEDGKNDSLPSESTSHRW. 616

655 PTYPSCR..DSETFSTF..LLDLFKLTIGMGDLEMLSSTKYPVVFIILLVTYIILTFVLL 710
617 .RGPACRPPDSSYNSLYSTCLELFKFTIGMGDLEFTENYDFKAVFIILLLAYVILTYILL 675

711 LNMLIALMGETVGQVSKESKHIWKLQWATTILDIERSFPVFLRKAFRSGEMVTVGKSSDG 770
676 LNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSFLKCMRKAFRSGKLLQVGYTPDG 735

771 TPDRRWCFRVDEVNWSHWNQNLGIINEDPGKNE.TYQYYGFSHTVGRLRRDRWS..SVVP 827
736 KDDYRWCFRVDEVNWTTWNTNVGIINEDPGNCEGVKRTLSFSLRSSRVSGRHWKNFALVP 795

828 RVVELN....KNSNPDEVVV.........PLDS..MGNPRCDGHQQGYPRKWRTDDAPL* 872
796 LLREASARDRQSAQPEEVYLRQFSGSLKPEDAEVFKSPAASGEK*.............. 840
```

FIG.6B

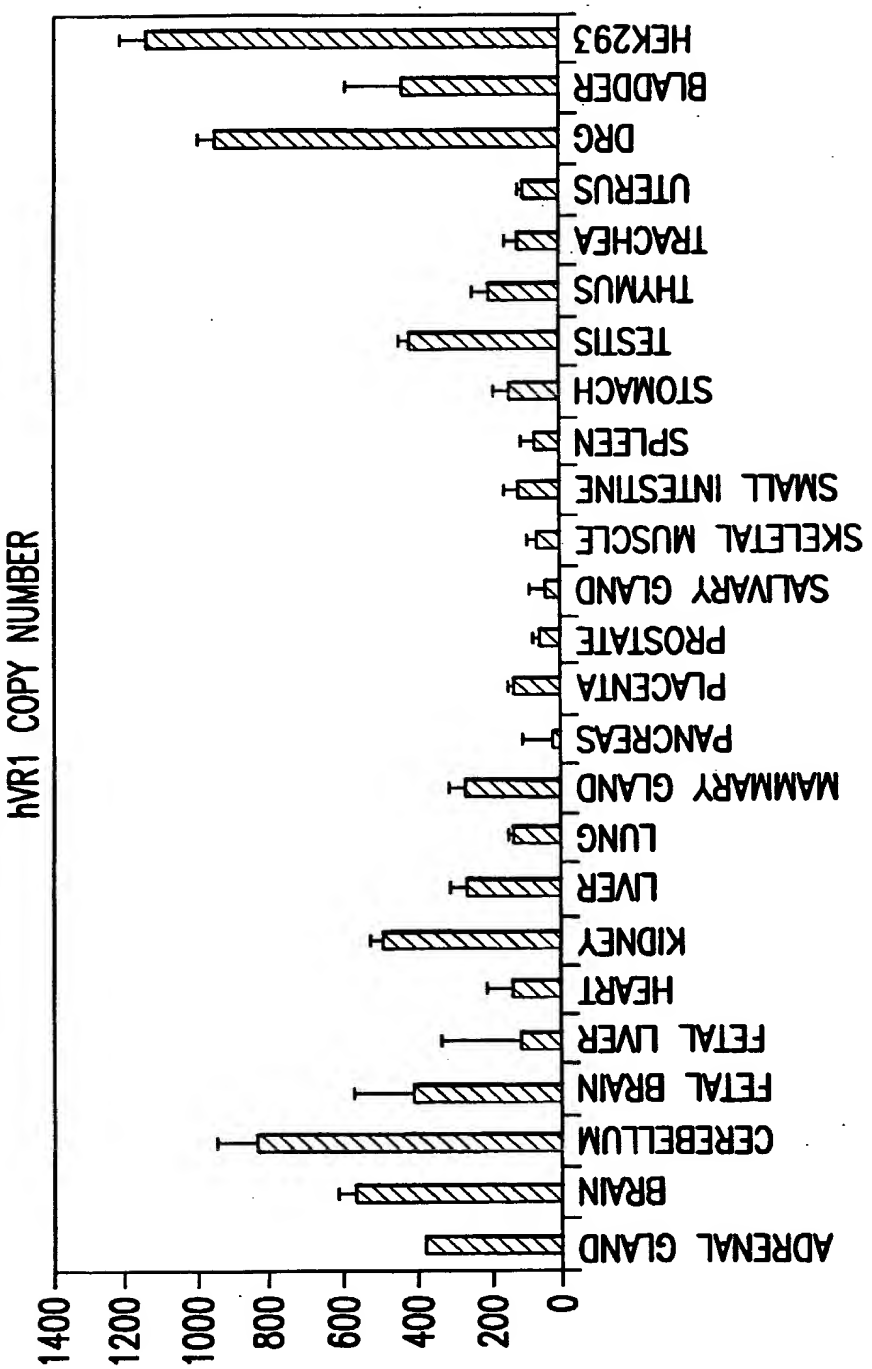
hVR1 COPY NUMBER
1400
1200
1000
800
600
400
200
0
ADRENAL GLAND
BRAIN
CEREBELLUM
FETAL BRAIN
FETAL LIVER
HEART
KIDNEY
LIVER
LUNG
MAMMARY GLAND
PANCREAS
PLACENTA
PROSTATE
SALIVARY GLAND
SKELETAL MUSCLE
SMALL INTESTINE
SPLEEN
STOMACH
TESTIS
THYMUS
TRACHEA
UTERUS
DRG
BLADDER
HEK293

FIG.7A

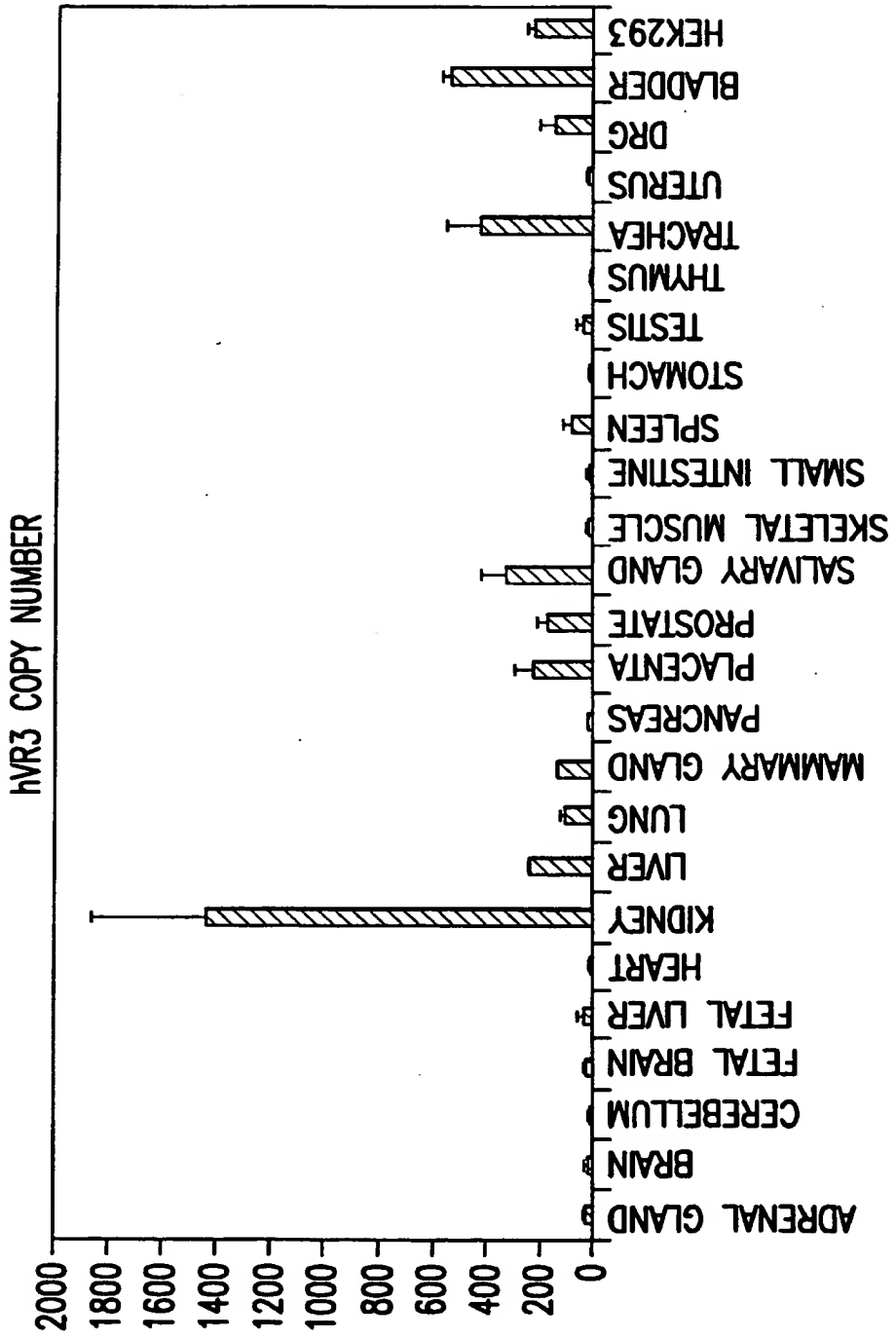
hVR3 COPY NUMBER
2000
1800
1600
1400
1200
1000
800
600
400
200
0
ADRENAL GLAND
BRAIN
CEREBELLUM
FETAL BRAIN
FETAL LIVER
HEART
KIDNEY
LIVER
LUNG
MAMMARY GLAND
PANCREAS
PLACENTA
PROSTATE
SALIVARY GLAND
SKELETAL MUSCLE
SMALL INTESTINE
SPLEEN
STOMACH
TESTIS
THYMUS
TRACHEA
UTERUS
DRG
BLADDER
HEK293

FIG.7B

```
gagccaccca ctgctccagc ctgccccagc tgttccctct gtctgtcctc tctgttttgc   60
agatggggaa actgaggctt aggtcgggga tctagacaat tgggatttaa acccagggac  120
tatccagccc caaagccctt cccaccacac caggtggcct gtcctggggc cagctctgca  180
cacagggcct ggtgcccccg gggtgcttgg gaagtggcag ggcagaggtg ggccctgtgg  240
ctgttctggc tcagcttcta aaacaagagc ctctgctggg ggcagagggg ccgtgaaccc  300
ctgaaatgtt aggcagatac cctgtgggag ctttgttctg ggatgctaag aaccgcttga  360
ggatttaagc tttgccactt tggctccgga gcaagggcag agggctgagc agtgcagacg  420
ggcctggggc aggc atg gcg gat tcc agc gaa ggc ccc cgc gcg ggg ccc     470
               Met Ala Asp Ser Ser Glu Gly Pro Arg Ala Gly Pro
                1               5                  10

ggg gag gtg gct gag ctc ccc ggg gat gag agt ggc acc cca ggt ggg     518
Gly Glu Val Ala Glu Leu Pro Gly Asp Glu Ser Gly Thr Pro Gly Gly
        15                  20                  25

gag gct ttt cct ctc tcc tcc ctg gcc aat ctg ttt gag ggg gag gat     566
Glu Ala Phe Pro Leu Ser Ser Leu Ala Asn Leu Phe Glu Gly Glu Asp
    30                  35                  40

ggc tcc ctt tcg ccc tca ccg gct gat gcc agt cgc cct gct ggc cca     614
Gly Ser Leu Ser Pro Ser Pro Ala Asp Ala Ser Arg Pro Ala Gly Pro
45                  50                  55                  60

ggc gat ggg cga cca aat ctg cgc atg aag ttc cag ggc gcc ttc cgc     662
Gly Asp Gly Arg Pro Asn Leu Arg Met Lys Phe Gln Gly Ala Phe Arg
                65                  70                  75

aag ggg gtg ccc aac ccc atc gat ctg ctg gag tgg acc cta tat gag     710
Lys Gly Val Pro Asn Pro Ile Asp Leu Leu Glu Ser Thr Leu Tyr Glu
            80                  85                  90

tcc tcg gtg gtg cct ggg ccc aag aaa gca ccc atg gac tca ctg ttt     758
Ser Ser Val Val Pro Gly Pro Lys Lys Ala Pro Met Asp Ser Leu Phe
        95                  100                 105

gac tac ggc acc tat cgt cac cac tcc agt gac aac aag agg tgg agg     806
Asp Tyr Gly Thr Tyr Arg His His Ser Ser Asp Asn Lys Arg Trp Arg
    110                 115                 120

aag aag atc ata gag aag cag ccg cag agc ccc aaa gcc cct gcc cct     854
Lys Lys Ile Ile Glu Lys Gln Pro GLn Ser Pro Lys Als Pro Ala Pro
125                 130                 130                 140

cag ccg ccc ccc atc ctc aaa gtc ttc aac cgg cct atc ctc ttt gac     902
Gln Pro Pro Pro Ile Leu Lys Val Phe Asn Arg Pro Ile Leu Phe Asp
                145                 150                 155

atc gtg tcc cgg ggc tcc act gct gac ctg gac ggg ctg ctc cca ttc     950
Ile Val Ser Arg Gly Ser Thr Ala Asp Leu Asp Gly Leu Leu Pro Phe
            160                 165                 170
```

FIG.8A

```
ttg ctg acc cac aag aaa cgc cta act gat gag gag ttt cga gag cca    998
Leu Leu Thr His Lys Lys Arg Leu Thr Asp Glu Glu Phe Arg Glu Pro
        175                 180                 185

tct acg ggg aag acc tgc ctg ccc agg gcc ttg ctg aac ctg agc aat   1046
Ser Thr Gly Lys Thr Cys Leu Pro Lys Ala Leu Leu Asn Leu Ser Asn
    190                 195                 200

ggc cgc aac gac acc atc cct gtg ctg ctg gac atc gcg gag cgc acc   1094
Gly Arg Asn Asp Thr Ile Pro Val Leu Leu Asp Ile Ala Glu Arg Thr
205                 210                 215                 220

ggc aac atg cgg gag ttc att aac tcg ccc ttc cgt gac atc tac tat   1142
Gly Asn Met Arg Glu Phe Ile Asn Ser Pro Phe Arg Asp Ile Tyr Tyr
                225                 230                 235

cga ggt cag aca gcc ctg cac atc gcc att gag cgt cgc tgc aaa cac   1190
Arg Gly Gln Thr Ala Leu His Ile Ala Ile Glu Arg Arg Cys Lys His
            240                 245                 250

tac gtg gaa ctt ctc gtg gcc cag gga gct gat gtc cac gcc cag gcc   1238
Tyr Val Glu Leu Leu Val Ala Gln Gly Ala Asp Val His Ala GLn Ala
        255                 260                 265

cgt ggg cgc ttc ttc cag ccc aag gat gag ggg ggc tac ttc tac ttt   1286
Arg Gly Arg Phe Phe GLn Pro Lys Asp Glu Gly Gly Tyr Phe Tyr Phe
    270                 275                 280

ggg gag ctg ccc ctg tcg ctg gct gcc tgc acc aac cag ccc cac att   1334
Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Asn Gln Pro His Ile
285                 290                 295                 300

gtc aac tac ctg acg aga aac ccc cac aag aag gcg gac atg cgg cgc   1382
Val Asn Tyr Leu Thr Glu Asn Pro His Lys Lys Ala Asp Met Arg Arg
                305                 310                 315

cag gac tcg cga ggc aac aca gtg ctg cat gcg ctg gtg gcc att gct   1430
Gln Asp Ser Arg Gly Asn Thr Val Leu His Ala Leu Val Ala Ile Ala
            320                 325                 330

gac aac acc cgt gag aac acc aag ttt gtt acc aag atg tac gac ctg   1478
Asp Asn Thr Arg Glu Asn Thr Lys Phe Val Thr Lys Met Tyr Asp Leu
        335                 340                 345

ctg ctg ctc aag tgt gcc cgc ctc ttc ccc gac agc aac ctg gag gcc   1526
Leu Leu Leu Lys Cys Ala Arg Leu Phe Pro Asp Ser Asn Leu Glu Ala
    350                 355                 360

gtg ctc aac aac gac ggc ctc tcg ccc ctc atg atg gct gcc aag acg   1574
Val Leu Asn Asn Asp Gly Leu Ser Pro Leu Met Met Ala Ala Lys Thr
365                 370                 375                 380
```

FIG.8B

```
ggc aag att ggg atc ttt cag cac atc atc cgg cgg gag gtg acg gat   1622
Gly Lys Ile Gly Ile Phe Gln His Ile Ile Arg Arg Glu Val Thr Asp
                385                 390                 395

gag gac aca cgg cac ctg tcc cgc aag ttc aag gac tgg gcc tat ggg   1670
Glu Asp Thr Arg His Leu Ser Arg Lys Phe Lys Asp Trp Ala Tyr Gly
            400                 405                 410

cca gtg tat tcc tcg ctt tat gac ctc tcc tcc ctg gac acg tgt ggg   1718
Pro Val Tyr Ser Ser Leu Tyr Asp Leu Ser Ser Leu Asp Thr Cys Gly
        415                 420                 425

gaa gag gcc tcc gtg ctg gag atc ctg gtg tac aac agc aag att gag   1766
Glu Glu Ala Ser Val Leu Glu Ile Leu Val Tyr Asn Ser Lys Ile Glu
    430                 435                 440

aac cgc cac gag atg ctg gct gtg gag ccc atc aat gaa ctg ctg cgg   1814
Asn Arg His Glu Met Leu Ala Val Glu Pro Ile Asn Glu Leu Leu Arg
445                 450                 455                 460

gac aag tgg cgc aag ttc ggg gcc gtc tcc ttc tac atc aac gtg gtc   1862
Asp Lys Trp Arg Lys Phe Gly Ala Val Ser Phe Tyr Ile Asn Val Val
                465                 470                 475

tcc tac ctg tgt gcc atg gtc atc ttc act ctc acc gcc tac tac cag   1910
Ser Tyr Leu Cys Ala Met Val Ile Phe Thr Leu Thr Ala Tyr Tyr Gln
            480                 485                 490

ccg ctg gag ggc aca ccg ccg tac cct tac cgc acc acg gtg gac tac   1958
Pro Leu Glu Gly Thr Pro Pro Tyr Pro Tyr Arg Thr Thr Val Asp Tyr
        495                 500                 505

ctg cgg ctg gct ggc gag gtc att acg ctc ttc act ggg gtc ctg ttc   2006
Leu Arg Leu Ala Gly Glu Val Ile Thr Leu Phe Thr Gly Val Leu Phe
    510                 515                 520

ttc ttc acc aac atc aaa gac ttg ttc atg aag aaa tgc cct gga gtg   2054
Phe Phe Thr Asn Ile Lys Asp Leu Phe Met Lys Lys Cys Pro Gly Val
525                 530                 535                 540

aat tct ctc ttc att gat ggc tcc ttc cag ctg ctc tac ttc atc tac   2102
Asn Ser Leu Phe Ile Asp Gly Ser Phe Gln Leu Leu Tyr Phe Ile Tyr
                545                 550                 555

tct gtc ctg gtg atc gtc tca gca gcc ctc tac ctg gca ggg atc gag   2150
Ser Val Leu Val Ile Val Ser Ala Ala Leu Tyr Leu Ala Gly Ile Glu
            560                 565                 570

gcc tac ctg gcc gtg atg gtc ttt gcc ctg gtc ctg ggc tgg atg aat   2198
Ala Tyr Leu Ala Val Met Val Phe Ala Leu Val Leu Gly Trp Met Asn
        575                 580                 585
```

FIG.8C

```
gcc ctt tac ttc acc cgt ggg ctg aag ctg acg ggg acc tat agc atc   2246
Ala Leu Tyr Phe THr Arg Gly Leu Lys Leu Thr Gly Thr Tyr Ser Ile
    590                 595                 600

atg atc cag aag att ctc ttc aag gac ctt ttc cga ttc ctg ctc gtc   2294
Met Ile Gly Lys Ile Leu Phe Lys Asp Leu Phe Arg Phe Leu Leu Val
605                 610                 615                 620

tac ttg ctc ttc atg atc ggc tac gct tca gcc ctg gtc tcc ctc ctg   2342
Tyr Leu Leu Phe Met Ile Gly Tyr Ala Ser Ala Leu Val Ser Leu Leu
                625                 630                 635

aac ccg tgt gcc aac atg aag gtg tgc aat gag gac cag acc aag tgc   2390
Asn Pro Cys Ala Asn Mey Lys Val Cys Asn Glu Asp Gln Thr Asn Cys
            640                 645                 650

aca gtg ccc act tac ccc tcg tgc cgt gac agc gag acc ttc agc acc   2438
Thr Val Pro Thr Tyr Pro Ser Cys Arg Asp Ser Glu Thr Phe Ser Thr
        655                 660                 665

ttc ctc ctg gac ctg ttt aag ctg acc atc ggc atg ggc gac ctg gag   2486
Phe Leu Leu Asp Lus Phe Lys Leu Thr Ile Gly Met Gly Asp Leu Glu
    670                 675                 680

atg ctg agc agc acc aag tac ccc gtg gtc ttc atc atc ctg ctg gtg   2534
Met Leu Ser Ser Thr Lys Tyr Pro Val Val Phe Ilr Ile Leu Leu Val
685                 690                 695                 700

acc tac atc atc ctc acc ttt gtg ctg ctc ctc aac atg ctc att gcc   2582
Thr Tyr Ile Ile Leu Thr Phe Val Leu Leu Leu Asn Met Leu Ile Ala
                705                 710                 715

ctc atg ggc gag aca gtg ggc cag gtc tcc aag gag agc aag cac atc   2630
Leu Met Gly Glu Thr Val Gly Gln Val Ser Lys Glu Ser Lys His Ile
            720                 725                 730

tgg aag ctg cag tgg gcc acc acc atc ctg gac att gag cgc tcc ttc   2678
Trp Lys Leu Gln Trp Ala Thr Thr Ile Leu Asp Ile Glu Arg Ser Phe
        735                 740                 745

ccc gta ttc ctg agg aag gcc ttc cgc tct ggg gag atg gtc acc gtg   2726
Pro Val Phe Leu Arg Lys Ala Phe Arg Ser Gly Glu Met Val Thr Val
    750                 755                 760

ggc aag agc tcg gac ggc act cct gac cgc agg tgg tgc ttc agg gtg   2774
Gly Lys Ser Ser Asp Gly Thr Pro Asp Arg Arg trp Cys Phe Arg Val
765                 770                 775                 780

gat gag gtg aac tgg tct cac tgg aac cag aac ttg ggc atc atc aac   2822
Asp Glu Val Asn Trp Ser His Trp Asn Gln Asn Leu Gly Ile Ile Asn
                785                 790                 795
```

FIG.8D

```
gag gac ccg ggc aag aat gag acc tac cag tat tat ggc ttc tcg cat   2870
Glu Asp Pro Gly Lys Asn Glu Thr Tyr Gln Tyr Tyr Gly Phe Ser His
            800                 805                 810

acc gtg ggc cgc ctc cgc agg gat cgc tgg tcc tcg gtg gta ccc cgc   2918
Thr Val Gly Arg Leu Arg Arg Asp Arg Trp Ser Ser Val Val Pro Arg
        815                 820                 825

gtg gtg gaa ctg aac aag aac tcg aac ccg gac gag gtg gtg gtg cct   2966
Val Val Glu Leu Asn Lys Asn Ser Asn Pro Asp Glu Val Val Val Pro
    830                 835                 840

ctg gac agc atg ggg aac ccc cgc tgc gat ggc cac cag cag ggt tac   3014
Leu Asp Ser Met Gly Asn Pro Arg Cys Asp Gly His Gln Gln Gly Tyr
845                 850                 855                 860

ccc cgc aag tgg agg act gat gac gcc ccg ctc tag ggactgcagc        3060
Pro Arg Lys Trp Arg Thr Asp Asp Ala Pro Leu  *
                865                 870

ccagccccag cttctctgcc cactcatttc tagtccagcc gcatttcagc agtgccttct 3120
ggggtgtccc cccacaccct gctttggccc cagaggcgag ggaccagtgg aggtgccagg 3180
gaggccccag gaccctgtgg tcccctggct ctgcctcccc accctggggt gggggctccc 3240
ggccacctgt cttgctccta tggagtcaca taagccaacg ccagagcccc tccacctcag 3300
gccccagccc ctgcctctcc attatttatt tgctctgctc tcaggaagcg acgtgacccc 3360
tgccccagct ggaacctggc agaggcctta ggaccccgtt ccaagtgcac tgcccggcca 3420
agccccagcc tcagcctgcg cctgagctgc atgcgccacc atttttggca gcgtggcagc 3480
tttgcaaggg gctggggccc tcggcgtggg gccatgcctt ctgtgtgttc tgtagtgtct 3540
gggatttgcc ggtgctcaat aaatgtttat tcattgaaaa aaaaaaaaaa aaaaaaaaa  3599
```

FIG.8E

SEQUENCE LISTING

<110> Abbott Laboratories
Masters, Jeffrey N.
Vos, Melissa H.

<120> Human Vanilloid Receptor Gene

<130> 6430.US.P1

<140> US 09/438,997
<141> 1999-11-12

<150> US 09/191,139
<151> 1998-11-13

<160> 16

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 3599
<212> DNA
<213> Homo sapiens

<400> 1
gagccaccca ctgctccagc ctgccccagc tgttccctct gtctgtcctc tctgttttgc 60
agatggggaa actgaggctt aggtcgggga tctagacaat tgggatttaa acccagggac 120
tatccagccc caaagccctt cccaccacac caggtggcct gtcctggggc cagctctgca 180
cacagggcct ggtgcccccg gggtgcttgg gaagtggcag ggcagaggtg ggccctgtgg 240
ctgttctggc tcagcttcta aaacaagagc ctctgctggg ggcagagggg ccgtgaaccc 300
ctgaaatgtt aggcagatac cctgtgggag ctttgttctg ggatgctaag aaccgcttga 360
ggatttaagc tttgccactt tggctccgga gcaagggcag agggctgagc agtgcagacg 420
ggcctggggc aggcatggcg gattccagcg aaggcccccg cgcggggccc ggggaggtgg 480
ctgagctccc cggggatgag agtggcaccc caggtgggga ggcttttcct ctctcctccc 540
tggccaatct gtttgagggg gaggatggct ccctttcgcc ctcaccggct gatgccagtc 600
gccctgctgg cccaggcgat gggcgaccaa atctgcgcat gaagttccag ggcgccttcc 660
gcaagggggt gcccaacccc atcgatctgc tggagtccac cctatatgag tcctcggtgg 720
tgcctgggcc caagaaagca cccatggact cactgtttga ctacggcacc tatcgtcacc 780
actccagtga caacaagagg tggaggaaga agatcataga gaagcagccg cagagcccca 840
aagcccctgc ccctcagccg ccccccatcc tcaaagtctt caaccggcct atcctctttg 900
acatcgtgtc ccggggctcc actgctgacc tggacgggct gctcccattc ttgctgaccc 960
acaagaaacg cctaactgat gaggagtttc gagagccatc tacggggaag acctgcctgc 1020
ccaaggcctt gctgaacctg agcaatggcc gcaacgacac catcctgtg ctgctggaca 1080
tcgcggagcg caccggcaac atgcgggagt tcattaactc gcccttccgt gacatctact 1140
atcgaggtca gacagccctg cacatcgcca ttgagcgtcg ctgcaaacac tacgtggaac 1200
ttctcgtggc ccagggagct gatgtccacg cccaggcccg tgggcgcttc ttccagccca 1260
aggatgaggg gggctacttc tactttgggg agctgcccct gtcgctggct gcctgcacca 1320
accagcccca cattgtcaac tacctgacgg agaaccccca caagaaggcg gacatgcggc 1380
gccaggactc gcgaggcaac acagtgctgc atgcgctggt ggccattgct gacaacaccc 1440
gtgagaacac caagtttgtt accaagatgt acgacctgct gctgctcaag tgtgcccgcc 1500
tcttccccga cagcaacctg gaggccgtgc tcaacaacga cggcctctcg cccctcatga 1560
tggctgccaa gacgggcaag attgggatct ttcagcacat catccggcgg gaggtgacgg 1620
atgaggacac acggcacctg tcccgcaagt tcaaggactg ggcctatggg ccagtgtatt 1680
cctcgcttta tgacctctcc tccctggaca cgtgtgggga agaggcctcc gtgctggaga 1740

tcctggtgta caacagcaag attgagaacc gccacgagat gctggctgtg gagcccatca 1800
atgaactgct gcgggacaag tggcgcaagt tcggggccgt ctccttctac atcaacgtgg 1860
tctcctacct gtgtgccatg gtcatcttca ctctcaccgc ctactaccag ccgctggagg 1920
gcacaccgcc gtaccttac cgcaccacgg tggactacct gcggctggct ggcgaggtca 1980
ttacgctctt cactggggtc ctgttcttct tcaccaacat caaagacttg ttcatgaaga 2040
aatgccctgg agtgaattct ctcttcattg atggctcctt ccagctgctc tacttcatct 2100
actctgtcct ggtgatcgtc tcagcagccc tctacctggc agggatcgag gcctacctgg 2160
ccgtgatggt ctttgccctg gtcctgggct ggatgaatgc cctttacttc acccgtgggc 2220
tgaagctgac ggggacctat agcatcatga tccagaagat tctcttcaag gacctttcc 2280
gattcctgct cgtctacttg ctcttcatga tcggctacgc ttcagccctg gtctccctcc 2340
tgaacccgtg tgccaacatg aaggtgtgca atgaggacca gaccaactgc acagtgccca 2400
cttacccctc gtgccgtgac agcgagacct tcagcacctt cctcctggac ctgtttaagc 2460
tgaccatcgg catgggcgac ctggagatgc tgagcagcac caagtacccc gtggtcttca 2520
tcatcctgct ggtgacctac atcatcctca cctttgtgct gctcctcaac atgctcattg 2580
ccctcatggg cgagacagtg ggccaggtct ccaaggagag caagcacatc tggaagctgc 2640
agtgggccac caccatcctg gacattgagc gctccttccc cgtattcctg aggaaggcct 2700
tccgctctgg ggagatggtc accgtgggca agagctcgga cggcactcct gaccgcaggt 2760
ggtgcttcag ggtggatgag gtgaactggt ctcactggaa ccagaacttg ggcatcatca 2820
acgaggaccc gggcaagaat gagacctacc agtattatgg cttctcgcat accgtgggcc 2880
gcctccgcag ggatcgctgg tcctcggtgg taccccgcgt ggtggaactg aacaagaact 2940
cgaacccgga cgaggtggtg gtgcctctgg acagcatggg gaacccccgc tgcgatggcc 3000
accagcaggg ttacccccgc aagtggagga ctgatgacgc cccgctctag ggactgcagc 3060
ccagccccag cttctctgcc cactcatttc tagtccagcc gcatttcagc agtgccttct 3120
ggggtgtccc cccacaccct gctttggccc cagaggcgag ggaccagtgg aggtgccagg 3180
gaggccccag gaccctgtgg tcccctggct ctgcctcccc accctggggt gggggctccc 3240
ggccacctgt cttgctccta tggagtcaca taagccaacg ccagagcccc tccacctcag 3300
gccccagccc ctgcctctcc attatttatt tgctctgctc tcaggaagcg acgtgacccc 3360
tgccccagct ggaacctggc agaggcctta ggaccccgtt ccaagtgcac tgcccggcca 3420
agccccagcc tcagcctgcg cctgagctgc atgcgccacc atttttggca gcgtggcagc 3480
tttgcaaggg gctggggccc tcggcgtggg gccatgcctt ctgtgtgttc tgtagtgtct 3540
gggatttgcc ggtgctcaat aaatgtttat tcattgaaaa aaaaaaaaaa aaaaaaaaa 3599

<210> 2
<211> 2687
<212> DNA
<213> Rattus norvegicus

<400> 2
gctggttgca aattgggcca cagaggatct ggaaaggatg gaacaacggg ctagcttaga 60
ctcagaggag tctgagtccc caccccaaga gaactcctgc ctggaccctc cagacagaga 120
ccctaactgc aagccacctc cagtcaagcc ccacatcttc actaccagga gtcgtacccg 180
gctttttggg aagggtgact cggaggaggc ctctcccctg gactgccctt atgaggaagg 240
cgggctggct tcctgcccta tcatcactgt cagctctgtt ctaactatcc agaggcctgg 300
ggatggacct gccagtgtca ggccgtcatc ccaggactcc gtctccgctg gtgagaagcc 360
cccgaggctc tatgatcgca ggagcatctt cgatgctgtg gctcagagta actgccagga 420
gctggagagc ctgctgccct tcctgcagag gagcaagaag cgcctgactg acagcgagtt 480
caaagaccca gagacaggaa agacctgtct gctaaaagcc atgctcaatc tgcacaatgg 540
gcagaatgac accatcgctc tgctcctgga cgttgcccgg aagacagaca gcctgaagca 600
gtttgtcaat gccagctaca cagacagcta ctacaagggc cagacagcac tgcacattgc 660
cattgaacgg cggaacatga cgctggtgac cctcttggtg gagaatggag cagatgtcca 720
ggctgcggct aacggggact tcttcaagaa aaccaaaggg aggcctggct tctactttgg 780
tgagctgccc ctgtccctgg ctgcgtgcac caaccagctg gccattgtga agttcctgct 840
gcagaactcc tggcagcctg cagacatcag cgcccgggac tcagtgggca acacggtgct 900
tcatgccctg gtggaggtgg cagataacac agttgacaac accaagttcg tgacaagcat 960
gtacaacgag atcttgatcc tgggggccaa actccacccc acgctgaagc tggaagagat 1020
caccaacagg aaggggctca cgccactggc tctggctgct agcagtggga agatcggggt 1080
cttggcctac attctccaga gggagatcca tgaacccgag tgccgacacc tatccaggaa 1140

gttcaccgaa tgggcctatg ggccagtgca ctcctccctt tatgacctgt cctgcattga 1200
cacctgtgaa aagaactcgg ttctggaggt gatcgcttac agcagcagtg agacccctaa 1260
ccgtcatgac atgcttctcg tggaaccctt gaaccgactc ctacaggaca agtgggacag 1320
atttgtcaag cgcatcttct acttcaactt cttcgtctac tgcttgtata tgatcatctt 1380
caccgcggct gcctactatc ggcctgtgga aggcttgccc ccctataagc tgaaaaacac 1440
cgttggggac tatttccgag tcaccggaga gatcttgtct gtgtcaggag gagtctactt 1500
cttcttccga gggattcaat atttcctgca gaggcgacca tccctcaaga gtttgtttgt 1560
ggacagctac agtgagatac ttttctttgt acagtcgctg ttcatgctgg tgtctgtggt 1620
actgtacttc agccaacgca aggagtatgt ggcttccatg gtgttctccc tggccatggg 1680
ctggaccaac atgctctact atacccgagg attccagcag atgggcatct atgctgtcat 1740
gattgagaag atgatcctca gagacctgtg ccggtttatg ttcgtctacc tcgtgttctt 1800
gtttggattt tccacagctg tggtgacact gattgaggat gggaagaata actctctgcc 1860
tatggagtcc acaccacaca agtgccgggg gtctgcctgc aagccaggta actcttacaa 1920
cagcctgtat tccacatgtc tggagctgtt caagttcacc atcggcatgg gcgacctgga 1980
gttcactgag aactacgact tcaaggctgt cttcatcatc ctgttactgg cctatgtgat 2040
tctcacctac atccttctgc tcaacatgct cattgctctc atgggtgaga ccgtcaacaa 2100
gattgcacaa gagagcaaga acatctggaa gctgcagaga gccatcacca tcctggatac 2160
agagaagagc ttcctgaagt gcatgaggaa ggccttccgc tctggcaagc tgctgcaggt 2220
ggggttcact cctgacggca aggatgacta ccggtggtgt ttcagggtgg acgaggtaaa 2280
ctggactacc tggaacacca atgtgggtat catcaacgag gacccaggca actgtgaggg 2340
cgtcaagcgc accctgagct tctccctgag gtcaggccga gtttcaggga gaaactggaa 2400
gaactttgcc ctggttcccc ttctgaggga tgcaagcact cgagatagac atgccaccca 2460
gcaggaagaa gttcaactga agcattatac gggatccctt aagccagagg atgctgaggt 2520
tttcaaggat tccatggtcc caggggagaa ataatggaca ctatgcaggg atcaatgcgg 2580
ggtctttggg tggtctgctt agggaaccag cagggttgac gttatctggg tccactctgt 2640
gcctgcctag gcacattcct aggacttcgg cgggcctgct gtgggaa 2687

<210> 3
<211> 871
<212> PRT
<213> Homo sapiens

<400> 3

```
Met Ala Asp Ser Ser Glu Gly Pro Arg Ala Gly Pro Gly Glu Val Ala
 1               5                  10                  15
Glu Leu Pro Gly Asp Glu Ser Gly Thr Pro Gly Gly Glu Ala Phe Pro
            20                  25                  30
Leu Ser Ser Leu Ala Asn Leu Phe Glu Gly Glu Asp Gly Ser Leu Ser
        35                  40                  45
Pro Ser Pro Ala Asp Ala Ser Arg Pro Ala Gly Pro Gly Asp Gly Arg
    50                  55                  60
Pro Asn Leu Arg Met Lys Phe Gln Gly Ala Phe Arg Lys Gly Val Pro
65                  70                  75                  80
Asn Pro Ile Asp Leu Leu Glu Ser Thr Leu Tyr Glu Ser Ser Val Val
                85                  90                  95
Pro Gly Pro Lys Lys Ala Pro Met Asp Ser Leu Phe Asp Tyr Gly Thr
            100                 105                 110
Tyr Arg His His Ser Ser Asp Asn Lys Arg Trp Arg Lys Lys Ile Ile
        115                 120                 125
Glu Lys Gln Pro Gln Ser Pro Lys Ala Pro Ala Pro Gln Pro Pro Pro
    130                 135                 140
Ile Leu Lys Val Phe Asn Arg Pro Ile Leu Phe Asp Ile Val Ser Arg
145                 150                 155                 160
Gly Ser Thr Ala Asp Leu Asp Gly Leu Leu Pro Phe Leu Leu Thr His
                165                 170                 175
Lys Lys Arg Leu Thr Asp Glu Glu Phe Arg Glu Pro Ser Thr Gly Lys
            180                 185                 190
```

```
Thr Cys Leu Pro Lys Ala Leu Leu Asn Leu Ser Asn Gly Arg Asn Asp
        195                 200                 205
Thr Ile Pro Val Leu Leu Asp Ile Ala Glu Arg Thr Gly Asn Met Arg
    210                 215                 220
Glu Phe Ile Asn Ser Pro Phe Arg Asp Ile Tyr Tyr Arg Gly Gln Thr
225                 230                 235                 240
Ala Leu His Ile Ala Ile Glu Arg Arg Cys Lys His Tyr Val Glu Leu
                245                 250                 255
Leu Val Ala Gln Gly Ala Asp Val His Ala Gln Ala Arg Gly Arg Phe
            260                 265                 270
Phe Gln Pro Lys Asp Glu Gly Gly Tyr Phe Tyr Phe Gly Glu Leu Pro
        275                 280                 285
Leu Ser Leu Ala Ala Cys Thr Asn Gln Pro His Ile Val Asn Tyr Leu
    290                 295                 300
Thr Glu Asn Pro His Lys Lys Ala Asp Met Arg Arg Gln Asp Ser Arg
305                 310                 315                 320
Gly Asn Thr Val Leu His Ala Leu Val Ala Ile Ala Asp Asn Thr Arg
                325                 330                 335
Glu Asn Thr Lys Phe Val Thr Lys Met Tyr Asp Leu Leu Leu Leu Lys
            340                 345                 350
Cys Ala Arg Leu Phe Pro Asp Ser Asn Leu Glu Ala Val Leu Asn Asn
        355                 360                 365
Asp Gly Leu Ser Pro Leu Met Met Ala Ala Lys Thr Gly Lys Ile Gly
    370                 375                 380
Ile Phe Gln His Ile Ile Arg Arg Glu Val Thr Asp Glu Asp Thr Arg
385                 390                 395                 400
His Leu Ser Arg Lys Phe Lys Asp Trp Ala Tyr Gly Pro Val Tyr Ser
                405                 410                 415
Ser Leu Tyr Asp Leu Ser Ser Leu Asp Thr Cys Gly Glu Glu Ala Ser
            420                 425                 430
Val Leu Glu Ile Leu Val Tyr Asn Ser Lys Ile Glu Asn Arg His Glu
        435                 440                 445
Met Leu Ala Val Glu Pro Ile Asn Glu Leu Leu Arg Asp Lys Trp Arg
    450                 455                 460
Lys Phe Gly Ala Val Ser Phe Tyr Ile Asn Val Val Ser Tyr Leu Cys
465                 470                 475                 480
Ala Met Val Ile Phe Thr Leu Thr Ala Tyr Tyr Gln Pro Leu Glu Gly
                485                 490                 495
Thr Pro Pro Tyr Pro Tyr Arg Thr Thr Val Asp Tyr Leu Arg Leu Ala
            500                 505                 510
Gly Glu Val Ile Thr Leu Phe Thr Gly Val Leu Phe Phe Phe Thr Asn
        515                 520                 525
Ile Lys Asp Leu Phe Met Lys Lys Cys Pro Gly Val Asn Ser Leu Phe
    530                 535                 540
Ile Asp Gly Ser Phe Gln Leu Leu Tyr Phe Ile Tyr Ser Val Leu Val
545                 550                 555                 560
Ile Val Ser Ala Ala Leu Tyr Leu Ala Gly Ile Glu Ala Tyr Leu Ala
                565                 570                 575
Val Met Val Phe Ala Leu Val Leu Gly Trp Met Asn Ala Leu Tyr Phe
            580                 585                 590
Thr Arg Gly Leu Lys Leu Thr Gly Thr Tyr Ser Ile Met Ile Gln Lys
        595                 600                 605
Ile Leu Phe Lys Asp Leu Phe Arg Phe Leu Leu Val Tyr Leu Leu Phe
    610                 615                 620
Met Ile Gly Tyr Ala Ser Ala Leu Val Ser Leu Leu Asn Pro Cys Ala
625                 630                 635                 640
Asn Met Lys Val Cys Asn Glu Asp Gln Thr Asn Cys Thr Val Pro Thr
```

645 650 655
Tyr Pro Ser Cys Arg Asp Ser Glu Thr Phe Ser Thr Phe Leu Leu Asp
660 665 670
Leu Phe Lys Leu Thr Ile Gly Met Gly Asp Leu Glu Met Leu Ser Ser
675 680 685
Thr Lys Tyr Pro Val Val Phe Ile Ile Leu Leu Val Thr Tyr Ile Ile
690 695 700
Leu Thr Phe Val Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu
705 710 715 720
Thr Val Gly Gln Val Ser Lys Glu Ser Lys His Ile Trp Lys Leu Gln
725 730 735
Trp Ala Thr Thr Ile Leu Asp Ile Glu Arg Ser Phe Pro Val Phe Leu
740 745 750
Arg Lys Ala Phe Arg Ser Gly Glu Met Val Thr Val Gly Lys Ser Ser
755 760 765
Asp Gly Thr Pro Asp Arg Arg Trp Cys Phe Arg Val Asp Glu Val Asn
770 775 780
Trp Ser His Trp Asn Gln Asn Leu Gly Ile Ile Asn Glu Asp Pro Gly
785 790 795 800
Lys Asn Glu Thr Tyr Gln Tyr Tyr Gly Phe Ser His Thr Val Gly Arg
805 810 815
Leu Arg Arg Asp Arg Trp Ser Ser Val Val Pro Arg Val Val Glu Leu
820 825 830
Asn Lys Asn Ser Asn Pro Asp Glu Val Val Val Pro Leu Asp Ser Met
835 840 845
Gly Asn Pro Arg Cys Asp Gly His Gln Gln Gly Tyr Pro Arg Lys Trp
850 855 860
Arg Thr Asp Asp Ala Pro Leu
865 870

<210> 4
<211> 838
<212> PRT
<213> Rattus norvegicus

<400> 4
Met Glu Gln Arg Ala Ser Leu Asp Ser Glu Glu Ser Glu Ser Pro Pro
1 5 10 15
Gln Glu Asn Ser Cys Leu Asp Pro Pro Asp Arg Asp Pro Asn Cys Lys
20 25 30
Pro Pro Pro Val Lys Pro His Ile Phe Thr Thr Arg Ser Arg Thr Arg
35 40 45
Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Ser Pro Leu Asp Cys Pro
50 55 60
Tyr Glu Glu Gly Gly Leu Ala Ser Cys Pro Ile Ile Thr Val Ser Ser
65 70 75 80
Val Leu Thr Ile Gln Arg Pro Gly Asp Gly Pro Ala Ser Val Arg Pro
85 90 95
Ser Ser Gln Asp Ser Val Ser Ala Gly Glu Lys Pro Pro Arg Leu Tyr
100 105 110
Asp Arg Arg Ser Ile Phe Asp Ala Val Ala Gln Ser Asn Cys Gln Glu
115 120 125
Leu Glu Ser Leu Leu Pro Phe Leu Gln Arg Ser Lys Lys Arg Leu Thr
130 135 140
Asp Ser Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu Lys
145 150 155 160

```
Ala Met Leu Asn Leu His Asn Gly Gln Asn Asp Thr Ile Ala Leu Leu
                165                 170                 175
Leu Asp Val Ala Arg Lys Thr Asp Ser Leu Lys Gln Phe Val Asn Ala
            180                 185                 190
Ser Tyr Thr Asp Ser Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile Ala
        195                 200                 205
Ile Glu Arg Arg Asn Met Thr Leu Val Thr Leu Leu Val Glu Asn Gly
    210                 215                 220
Ala Asp Val Gln Ala Ala Ala Asn Gly Asp Phe Phe Lys Lys Thr Lys
225                 230                 235                 240
Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala Ala
                245                 250                 255
Cys Thr Asn Gln Leu Ala Ile Val Lys Phe Leu Leu Gln Asn Ser Trp
            260                 265                 270
Gln Pro Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val Leu
        275                 280                 285
His Ala Leu Val Glu Val Ala Asp Asn Thr Val Asp Asn Thr Lys Phe
    290                 295                 300
Val Thr Ser Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu His
305                 310                 315                 320
Pro Thr Leu Lys Leu Glu Glu Ile Thr Asn Arg Lys Gly Leu Thr Pro
                325                 330                 335
Leu Ala Leu Ala Ala Ser Ser Gly Lys Ile Gly Val Leu Ala Tyr Ile
            340                 345                 350
Leu Gln Arg Glu Ile His Glu Pro Glu Cys Arg His Leu Ser Arg Lys
        355                 360                 365
Phe Thr Glu Trp Ala Tyr Gly Pro Val His Ser Ser Leu Tyr Asp Leu
    370                 375                 380
Ser Cys Ile Asp Thr Cys Glu Lys Asn Ser Val Leu Glu Val Ile Ala
385                 390                 395                 400
Tyr Ser Ser Ser Glu Thr Pro Asn Arg His Asp Met Leu Leu Val Glu
                405                 410                 415
Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp Arg Phe Val Lys Arg
            420                 425                 430
Ile Phe Tyr Phe Asn Phe Phe Val Tyr Cys Leu Tyr Met Ile Ile Phe
        435                 440                 445
Thr Ala Ala Ala Tyr Tyr Arg Pro Val Glu Gly Leu Pro Pro Tyr Lys
    450                 455                 460
Leu Lys Asn Thr Val Gly Asp Tyr Phe Arg Val Thr Gly Glu Ile Leu
465                 470                 475                 480
Ser Val Ser Gly Gly Val Tyr Phe Phe Phe Arg Gly Ile Gln Tyr Phe
                485                 490                 495
Leu Gln Arg Arg Pro Ser Leu Lys Ser Leu Phe Val Asp Ser Tyr Ser
            500                 505                 510
Glu Ile Leu Phe Phe Val Gln Ser Leu Phe Met Leu Val Ser Val Val
        515                 520                 525
Leu Tyr Phe Ser Gln Arg Lys Glu Tyr Val Ala Ser Met Val Phe Ser
    530                 535                 540
Leu Ala Met Gly Trp Thr Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln
545                 550                 555                 560
Gln Met Gly Ile Tyr Ala Val Met Ile Glu Lys Met Ile Leu Arg Asp
                565                 570                 575
Leu Cys Arg Phe Met Phe Val Tyr Leu Val Phe Leu Phe Gly Phe Ser
            580                 585                 590
Thr Ala Val Val Thr Leu Ile Glu Asp Gly Lys Asn Asn Ser Leu Pro
        595                 600                 605
Met Glu Ser Thr Pro His Lys Cys Arg Gly Ser Ala Cys Lys Pro Gly
```

```
    610                 615                 620
Asn Ser Tyr Asn Ser Leu Tyr Ser Thr Cys Leu Glu Leu Phe Lys Phe
625                 630                 635                 640
Thr Ile Gly Met Gly Asp Leu Glu Phe Thr Glu Asn Tyr Asp Phe Lys
                645                 650                 655
Ala Val Phe Ile Ile Leu Leu Leu Ala Tyr Val Ile Leu Thr Tyr Ile
            660                 665                 670
Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val Asn Lys
        675                 680                 685
Ile Ala Gln Glu Ser Lys Asn Ile Trp Lys Leu Gln Arg Ala Ile Thr
    690                 695                 700
Ile Leu Asp Thr Glu Lys Ser Phe Leu Lys Cys Met Arg Lys Ala Phe
705                 710                 715                 720
Arg Ser Gly Lys Leu Leu Gln Val Gly Phe Thr Pro Asp Gly Lys Asp
                725                 730                 735
Asp Tyr Arg Trp Cys Phe Arg Val Asp Glu Val Asn Trp Thr Thr Trp
            740                 745                 750
Asn Thr Asn Val Gly Ile Ile Asn Glu Asp Pro Gly Asn Cys Glu Gly
        755                 760                 765
Val Lys Arg Thr Leu Ser Phe Ser Leu Arg Ser Gly Arg Val Ser Gly
    770                 775                 780
Arg Asn Trp Lys Asn Phe Ala Leu Val Pro Leu Leu Arg Asp Ala Ser
785                 790                 795                 800
Thr Arg Asp Arg His Ala Thr Gln Gln Glu Glu Val Gln Leu Lys His
                805                 810                 815
Tyr Thr Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Asp Ser
            820                 825                 830
Met Val Pro Gly Glu Lys
        835
```

<210> 5
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 5
tttaaacgcg taggatgaag aaatggagc 29

<210> 6
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 6
tatattgtcg acgtcctcac ttctccccg 29

<210> 7
<211> 2712
<212> DNA
<213> Homo sapiens

<400> 7
gcaggttgca cactgggcca cagaggatcc agcaaggatg aagaaatgga gcagcacaga 60
cttgggggca gctgcggacc cactccaaaa ggacacctgc ccagaccccc tggatggaga 120
ccctaactcc aggccacctc cagccaagcc ccagctctcc acggccaaga gccgcacccg 180
gctctttggg aagggtgact cggaggaggc tttcccggtg gattgccctc acgaggaagg 240
tgagctggac tcctgcccga ccatcacagt cagccctgtt atcaccatcc agaggccagg 300
agacggcccc accggtgcca ggctgctgtc ccaggactct gtcgccgcca gcaccgagaa 360
gaccctcagg ctctatgatc gcaggagtat ctttgaagcc gttgctcaga ataactgcca 420
ggatctggag agcctgctgc tcttcctgca gaagagcaag aagcacctca cagacaacga 480
gttcaaagac cctgagacag ggaagacctg tctgctgaaa gccatgctca acctgcacga 540
cggacagaac accaccatcc ccctgctcct ggagatcgcg cggcaaacgg acagcctgaa 600
ggagcttgtc aacgccagct acacggacag ctactacaag ggccagacag cactgcacat 660
cgccatcgag agacgcaaca tggccctggt gaccctcctg gtggagaacg gagcagacgt 720
ccaggctgcg gcccatgggg acttctttaa gaaaaccaaa gggcggcctg gattctactt 780
cggtgaactg cccctgtccc tggccgcgtg caccaaccag ctgggcatcg tgaagttcct 840
gctgcagaac tcctggcaga cggccgacat cagcgccagg gactcggtgg gcaacacggt 900
gctgcacgcc ctggtggagg tggccgacaa cacggccgac aacacgaagt ttgtgacgag 960
catgtacaat gagattctga tcctgggggc caaactgcac ccgacgctga agctggagga 1020
gctcaccaac aagaagggaa tgacgccgct ggctctggca gctgggaccg ggaagatcgg 1080
ggtcttggcc tatattctcc agcgggagat ccaggagccc gagtgcaggc acctgtccag 1140
gaagttcacc gagtgggcct acgggcccgt gcactcctcg ctgtacgacc tgtcctgcat 1200
cgacacctgc gagaagaact cggtgctgga ggtgatcgcc tacagcagca gcgagacccc 1260
taatcgccac gacatgctct tggtggagcc gctgaaccga ctcctgcagg acaagtggga 1320
cagattcgtc aagcgcatct tctacttcaa cttcctggtc tactgcctgt acatgatcat 1380
cttcaccatg gctgcctact acaggcccgt ggatggcttg cctcccttta agatggaaaa 1440
aactggagac tatttccgag ttactggaga gatcctgtct gtgttaggag gagtctactt 1500
ctttttccga gggattcagt atttcctgca gaggcggccg tcgatgaaga ccctgtttgt 1560
ggacagctac agtgagatgc ttttctttct gcagtcactg ttcatgctgg ccaccgtggt 1620
gctgtacttc agccacctca aggagtatgt ggcttccatg gtattctccc tggccttggg 1680
ctggaccaac atgctctact acacccgcgg tttccagcag atgggcatct atgccgtcat 1740
gattgagaag atgatcctga gagacctgtg ccgtttcatg tttgtctaca tcgtcttctt 1800
gttcgggttt tccacagcgg tggtgacgct gattgaagac gggaagaatg actccctgcc 1860
gtctgagtcc acgtcgcaca ggtggcgggg gcctgcctgc aggcccccg atagctccta 1920
caacagcctg tactccacct gcctggagct gttcaagttc accatcggca tgggcgacct 1980
ggagttcact gagaactatg acttcaaggc tgtcttcatc atcctgctgc tggcctatgt 2040
aattctcacc tacatcctcc tgctcaacat gctcatcgcc ctcatgggtg agactgtcaa 2100
caagatcgca caggagagca agaacatctg gaagctgcag agagccatca ccatcctgga 2160
cacggagaag agcttcctta agtgcatgag gaaggccttc cgctcaggca agctgctgca 2220
ggtggggtac acacctgatg gcaaggacga ctaccggtgg tgcttcaggg tggacgaggt 2280
gaactggacc acctggaaca ccaacgtggg catcatcaac gaagacccgg gcaactgtga 2340
gggcgtcaag cgcaccctga gcttctccct gcggtcaagc agagtttcag gcagacactg 2400
gaagaacttt gccctggtcc ccctttaag agaggcaagt gctcgagata ggcagtctgc 2460
tcagcccgag gaagtttatc tgcgacagtt ttcagggtct ctgaagccag aggacgctga 2520
ggtcttcaag agtcctgccg cttccgggga gaagtgagga cgtcacgcag acagcactgt 2580
caacactggg ccttaggaga ccccgttgcc acgggggct tgctgaggga acaccagtgc 2640
tctgtcagca gcctggcctg gtctgtgcct gcccagcatg ttcccaaatc tgtgctggac 2700
aagctgtggg aa 2712

<210> 8
<211> 839
<212> PRT
<213> Homo sapiens

<400> 8
Met Lys Lys Trp Ser Ser Thr Asp Leu Gly Ala Ala Ala Asp Pro Leu
1 5 10 15

```
Gln Lys Asp Thr Cys Pro Asp Pro Leu Asp Gly Asp Pro Asn Ser Arg
            20                  25                  30
Pro Pro Pro Ala Lys Pro Gln Leu Ser Thr Ala Lys Ser Arg Thr Arg
        35                  40                  45
Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Phe Pro Val Asp Cys Pro
    50                  55                  60
His Glu Glu Gly Glu Leu Asp Ser Cys Pro Thr Ile Thr Val Ser Pro
65                  70                  75                  80
Val Ile Thr Ile Gln Arg Pro Gly Asp Gly Pro Thr Gly Ala Arg Leu
                85                  90                  95
Leu Ser Gln Asp Ser Val Ala Ala Ser Thr Glu Lys Thr Leu Arg Leu
            100                 105                 110
Tyr Asp Arg Arg Ser Ile Phe Glu Ala Val Ala Gln Asn Asn Cys Gln
        115                 120                 125
Asp Leu Glu Ser Leu Leu Leu Phe Leu Gln Lys Ser Lys Lys His Leu
    130                 135                 140
Thr Asp Asn Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu
145                 150                 155                 160
Lys Ala Met Leu Asn Leu His Asp Gly Gln Asn Thr Thr Ile Pro Leu
                165                 170                 175
Leu Leu Glu Ile Ala Arg Gln Thr Asp Ser Leu Lys Glu Leu Val Asn
            180                 185                 190
Ala Ser Tyr Thr Asp Ser Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile
        195                 200                 205
Ala Ile Glu Arg Arg Asn Met Ala Leu Val Thr Leu Leu Val Glu Asn
    210                 215                 220
Gly Ala Asp Val Gln Ala Ala Ala His Gly Asp Phe Phe Lys Lys Thr
225                 230                 235                 240
Lys Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala
                245                 250                 255
Ala Cys Thr Asn Gln Leu Gly Ile Val Lys Phe Leu Leu Gln Asn Ser
            260                 265                 270
Trp Gln Thr Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val
        275                 280                 285
Leu His Ala Leu Val Glu Val Ala Asp Asn Thr Ala Asp Asn Thr Lys
    290                 295                 300
Phe Val Thr Ser Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu
305                 310                 315                 320
His Pro Thr Leu Lys Leu Glu Glu Leu Thr Asn Lys Lys Gly Met Thr
                325                 330                 335
Pro Leu Ala Leu Ala Ala Gly Thr Gly Lys Ile Gly Val Leu Ala Tyr
            340                 345                 350
Ile Leu Gln Arg Glu Ile Gln Glu Pro Glu Cys Arg His Leu Ser Arg
        355                 360                 365
Lys Phe Thr Glu Trp Ala Tyr Gly Pro Val His Ser Ser Leu Tyr Asp
    370                 375                 380
Leu Ser Cys Ile Asp Thr Cys Glu Lys Asn Ser Val Leu Glu Val Ile
385                 390                 395                 400
Ala Tyr Ser Ser Ser Glu Thr Pro Asn Arg His Asp Met Leu Leu Val
                405                 410                 415
Glu Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp Arg Phe Val Lys
            420                 425                 430
Arg Ile Phe Tyr Phe Asn Phe Leu Val Tyr Cys Leu Tyr Met Ile Ile
        435                 440                 445
Phe Thr Met Ala Ala Tyr Tyr Arg Pro Val Asp Gly Leu Pro Pro Phe
    450                 455                 460
Lys Met Glu Lys Thr Gly Asp Tyr Phe Arg Val Thr Gly Glu Ile Leu
```

```
465                 470                 475                 480
Ser Val Leu Gly Gly Val Tyr Phe Phe Phe Arg Gly Ile Gln Tyr Phe
                485                 490                 495
Leu Gln Arg Arg Pro Ser Met Lys Thr Leu Phe Val Asp Ser Tyr Ser
            500                 505                 510
Glu Met Leu Phe Phe Leu Gln Ser Leu Phe Met Leu Ala Thr Val Val
        515                 520                 525
Leu Tyr Phe Ser His Leu Lys Glu Tyr Val Ala Ser Met Val Phe Ser
    530                 535                 540
Leu Ala Leu Gly Trp Thr Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln
545                 550                 555                 560
Gln Met Gly Ile Tyr Ala Val Met Ile Glu Lys Met Ile Leu Arg Asp
                565                 570                 575
Leu Cys Arg Phe Met Phe Val Tyr Ile Val Phe Leu Phe Gly Phe Ser
            580                 585                 590
Thr Ala Val Val Thr Leu Ile Glu Asp Gly Lys Asn Asp Ser Leu Pro
        595                 600                 605
Ser Glu Ser Thr Ser His Arg Trp Arg Gly Pro Ala Cys Arg Pro Pro
    610                 615                 620
Asp Ser Ser Tyr Asn Ser Leu Tyr Ser Thr Cys Leu Glu Leu Phe Lys
625                 630                 635                 640
Phe Thr Ile Gly Met Gly Asp Leu Glu Phe Thr Glu Asn Tyr Asp Phe
                645                 650                 655
Lys Ala Val Phe Ile Ile Leu Leu Leu Ala Tyr Val Ile Leu Thr Tyr
            660                 665                 670
Ile Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val Asn
        675                 680                 685
Lys Ile Ala Gln Glu Ser Lys Asn Ile Trp Lys Leu Gln Arg Ala Ile
    690                 695                 700
Thr Ile Leu Asp Thr Glu Lys Ser Phe Leu Lys Cys Met Arg Lys Ala
705                 710                 715                 720
Phe Arg Ser Gly Lys Leu Leu Gln Val Gly Tyr Thr Pro Asp Gly Lys
                725                 730                 735
Asp Asp Tyr Arg Trp Cys Phe Arg Val Asp Glu Val Asn Trp Thr Thr
            740                 745                 750
Trp Asn Thr Asn Val Gly Ile Ile Asn Glu Asp Pro Gly Asn Cys Glu
        755                 760                 765
Gly Val Lys Arg Thr Leu Ser Phe Ser Leu Arg Ser Ser Arg Val Ser
    770                 775                 780
Gly Arg His Trp Lys Asn Phe Ala Leu Val Pro Leu Leu Arg Glu Ala
785                 790                 795                 800
Ser Ala Arg Asp Arg Gln Ser Ala Gln Pro Glu Glu Val Tyr Leu Arg
                805                 810                 815
Gln Phe Ser Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Ser
            820                 825                 830
Pro Ala Ala Ser Gly Glu Lys
        835
```

<210> 9
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 9
tgcctggagc tgttcaagtt c 21

<210> 10
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 10
tgatgaagac agccttgaag tca 23

<210> 11
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 11
agttctcagt gaactccagg tcgcccat 28

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 12
attgagaacc gccacgagat 20

<210> 13
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 13
agacggcccc gaacttg 17

<210> 14
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 14
ccatcaatga actgctgcgg gaca 24

<210> 15
<211> 764
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> (670)...(670)
<223> Xaa = Unknown or other at position 670

<400> 15
Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
1 5 10 15
Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
20 25 30
Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
35 40 45
Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
50 55 60
Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
65 70 75 80
Asn Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
85 90 95
Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
100 105 110
Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
115 120 125
Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
130 135 140
Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
145 150 155 160
Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
165 170 175
Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
180 185 190
Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe
195 200 205
Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
210 215 220
Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
225 230 235 240
Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
245 250 255
Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
260 265 270
Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
275 280 285
Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
290 295 300
Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
305 310 315 320
Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
325 330 335
Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
340 345 350
Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His

```
        355                 360                 365
Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp
    370                 375                 380
Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile
385                 390                 395                 400
Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys
                405                 410                 415
Lys Gln Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu
            420                 425                 430
Leu Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val
        435                 440                 445
Gly Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser
    450                 455                 460
Phe Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu
465                 470                 475                 480
Thr Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu
                485                 490                 495
Pro Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr
            500                 505                 510
Tyr Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln
        515                 520                 525
Lys Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val
    530                 535                 540
Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala
545                 550                 555                 560
Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln
                565                 570                 575
Pro Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly
            580                 585                 590
Ile Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly
        595                 600                 605
Glu Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu
    610                 615                 620
Leu Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met
625                 630                 635                 640
Leu Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser
                645                 650                 655
Trp Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Xaa Met Glu
            660                 665                 670
Asn Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu
        675                 680                 685
Thr Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe
    690                 695                 700
Arg Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
705                 710                 715                 720
Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn
                725                 730                 735
Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu
            740                 745                 750
Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
        755                 760
```

<210> 16
<211> 244
<212> PRT
<213> Artificial Sequence

<220>
<223> Consensus

<400> 16

```
Arg Phe Val Leu Leu Leu Lys Leu Thr Asp Glu Thr Gly Lys Thr Cys
 1               5                  10                  15
Leu Lys Ala Leu Asn Leu Gly Asn Ile Leu Leu Asn Asp Tyr Tyr Gly
            20                  25                  30
Ala Leu His Ile Ala Ile Glu Arg Val Leu Leu Val Gly Ala Val Ala
        35                  40                  45
Ala Gly Phe Phe Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala Ala
    50                  55                  60
Cys Thr Gln Val Leu Asn Ala Asp Ser Gly Asn Thr Val Leu His Ala
65                  70                  75                  80
Leu Val Asp Asn Asn Val Thr Met Tyr Leu Ala Leu Pro Leu Glu Asn
                85                  90                  95
Pro Leu Ala Ala Gly Lys Ile Ile Arg Glu His Leu Ser Arg Lys Phe
            100                 105                 110
Trp Tyr Gly Pro Val Ser Leu Tyr Asp Leu Asp Cys Glu Ser Val Leu
        115                 120                 125
Glu Arg His Met Glu Pro Asn Leu Leu Lys Trp Phe Asn Met Ile Phe
    130                 135                 140
Thr Ala Tyr Pro Gly Gly Phe Phe Asp Leu Leu Tyr Val Leu Gly Trp
145                 150                 155                 160
Asn Leu Tyr Thr Arg Gly Gly Tyr Met Ile Lys Asp Leu Arg Phe Tyr
                165                 170                 175
Phe Gly Ala Val Leu Leu Leu Phe Lys Thr Ile Gly Met Gly Leu Leu
            180                 185                 190
Leu Tyr Leu Thr Leu Leu Leu Asn Met Leu Ile Ala Leu Met Glu Thr
        195                 200                 205
Val Ser Ile Trp Lys Leu Gln Ala Leu Glu Arg Lys Arg Gly Val Gly
    210                 215                 220
Asp Gly Asp Arg Trp Cys Phe Arg Val Glu Val Asn Trp Trp Glu Asp
225                 230                 235                 240
Pro Thr Val Leu
```